97526

STIC-Biotech/ChemLib

From: Rao, Manjunath N.
Sent: Thursday, June 26, 2003 11:20 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/762,767

Fr m: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10D 01
Phone: 306-5681

Date: 6-26-03

Please search the following as soon as possible for application with serial number **09/762,767**

1. **SEQ ID NO:1 and 3,** against all **commercial nucleic acid databases** including **issued patents database** and **pending application database** and provide **a print of all results.**

2. **SEQ ID NO: 2 and 4 ,** against all **commercial protein databases** including **issued patents database** and **pending application database** and provide **a print of all results.**

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Mail Box in 10D01
Crystal Mall 1, USPTO.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 12:42:07 ; Search time 36.3713 Seconds
(without alignments)
2527.899 Million cell updates/sec

Title: US-09-762-767A-2
Perfect score: 3593
Sequence: 1 MFCTKLKDLKITGECPFSLL.....QKKDVEDGNANFLGKASGID 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3593	100.0	690	21	AAY70475	Human cyclic nucle
2	3593	100.0	690	21	AAY51607	Human soluble guan
3	3593	100.0	690	23	ABG61940	Prostate cancer-as
4	3171.5	88.3	717	23	ABG61878	Prostate cancer-as
5	991.5	27.6	676	22	ABB58042	Drosophila melanog
6	789.5	22.0	619	21	AAY51608	Human soluble guan
7	743	20.7	787	22	ABB58149	Drosophila melanog
8	642.5	17.9	690	22	ABB60360	Drosophila melanog
9	462	12.9	1076	22	ABB59684	Drosophila melanog
10	461	12.8	1115	22	AAB83940	Amino acid sequenc
11	460	12.8	1025	14	AAR38863	GC-B. Rattus ratt
12	460	12.8	1047	12	AAR10399	Human Natriuretic
13	460	12.8	1047	12	AAR10867	NPRB(Pro655, Glu65
14	457.5	12.7	1029	14	AAR38862	GC-A. Rattus ratt
15	456.5	12.7	1061	23	AAU11280	Human atrionatriur
16	456.5	12.7	1061	23	AAU11281	Human atrionatriur
17	456.5	12.7	1075	22	ABB11783	Human ANP-A recept
18	456.5	12.7	1172	22	ABB72034	Drosophila melanog
19	453	12.6	1076	22	ABB71206	Drosophila melanog
20	445	12.4	975	22	ABB64661	Drosophila melanog
21	423.5	11.8	1525	22	ABB71674	Drosophila melanog
22	421.5	11.7	566	22	ABB61743	Drosophila melanog
23	408.5	11.4	887	22	ABG03067	Novel human diagno
24	408.5	11.4	1073	18	AAW32063	Human ST receptor
25	408.5	11.4	1073	19	AAW37371	Human ST receptor
26	408.5	11.4	1073	23	AAU08788	Human guanylin cyc
27	407.5	11.3	1075	14	AAR38861	GC-C. Rattus ratt
28	403.5	11.2	1151	22	AAB83939	Amino acid sequenc
29	398.5	11.1	949	22	AAB83941	Amino acid sequenc
30	397	11.0	669	22	ABB69903	Drosophila melanog
31	396	11.0	649	22	ABB69902	Drosophila melanog
32	356	9.9	273	22	AAG67400	Partial amino acid
33	325.5	9.1	1077	22	AAU01924	Human adenylate cy
34	322.5	9.0	887	23	AAE14741	Human adenylate cy
35	322.5	9.0	1086	22	AAE02938	Human adenylate cy
36	322.5	9.0	1086	23	AAE14742	Human adenylate cy
37	322	9.0	1064	21	AAB02008	Type IV adenylyl c
38	319.5	8.9	1099	21	AAB02011	Type VII adenylyl
39	319	8.9	1090	17	AAR94560	Rat adenylyl cycla
40	319	8.9	1090	21	AAB02006	Adenylyl cyclase t
41	312	8.7	1086	23	AAE17129	Human adenylyl and
42	310.5	8.6	1103	23	ABB81465	Human adenylate cy
43	303.5	8.4	855	21	AAB54359	Human pancreatic c
44	303.5	8.4	857	23	ABG64609	Human albumin fusi
45	303.5	8.4	857	23	AAE21639	Human gene 11 enco

ALIGNMENTS

RESULT 1
AAY70475
ID AAY70475 standard; Protein; 690 AA.
XX
AC AAY70475;
XX
DT 04-JUL-2000 (first entry)
XX
DE Human cyclic nucleotide-associated protein-3 (CNAP-3).
XX
KW Cyclic nucleotide-associated protein-3; CNAP-3; human; cytostatic;
KW anti-arteriosclerotic; hepatotropic; anti-leukaemic; anti-inflammatory;
KW immunomodulatory; anti-asthmatic; anti-anaemic; anti-diabetic; diagnosis;
KW anti-sclerotic; dermatological; neuroprotective; anti-epileptic; cancer;
KW anti-Alzheimer's; anti-Parkinsonian; cerebroprotective; ophthalmological;
KW anti-infertility; anti-allergic; vasotropic; immunosuppressive;
KW hypotensive; gene therapy; prevention; treatment; arteriosclerosis;
KW cell proliferative disorder; autoimmune/inflammatory; diabetes mellitus;
KW neurological; vision; reproductive; smooth muscle.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 65
FT /note= "Potential phosphorylation site"
FT Modified-site 211
FT /note= "Potential phosphorylation site"
FT Modified-site 361
FT /note= "Potential phosphorylation site"
FT Modified-site 450
FT /note= "Potential phosphorylation site"
FT Modified-site 478
FT /note= "Potential phosphorylation site"

```
FT   Modified-site   29
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   73
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   117
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   135
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   158
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   229
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   511
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   527
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   638
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   687
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   39
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   64
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   105
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   261
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   328
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   356
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   360
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   394
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   622
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   635
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   27
FT                   /note= "N-glycosylated"
FT   Modified-site   332
FT                   /note= "N-glycosylated"
FT   Modified-site   479
FT                   /note= "N-glycosylated"
FT   Modified-site   600
FT                   /note= "N-glycosylated"
FT   Region          472..661
FT                   /note= "Guanylate cyclase signature sequence
FT                   identified by PFAM analysis; CNAP-3 shares 89%
FT                   identity with human soluble guanylate cyclase
FT                   large subunit"
XX
PN   WO200014248-A1.
XX
PD   16-MAR-2000.
XX
PF   03-SEP-1999;   99WO-US20287.
XX
PR   04-SEP-1998;   98US-0148904.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Yue H,  Guegler KJ,  Corley NC,  Patterson C,  Tang YT;
XX
DR   WPI; 2000-256994/22.
DR   N-PSDB; AAZ51684.
XX
PT   Isolated cyclic nucleotide associated proteins useful for preventing,
PT   diagnosing and treating cell proliferative, autoimmune/inflammatory,
PT   neurological, vision, reproductive and smooth muscle disorders -
XX
PS   Disclosure; Page 67-69; 78pp; English.
XX
CC   The present sequence is a human cyclic nucleotide
CC   associated protein-3 (CNAP-3), identified in Incyte clone 159278,
CC   that is isolated from ADENINB01 cDNA library. It is expressed in
CC   nervous, reproductive, cardiovascular and developmental tissues.
CC   CNAP sequences may be used for prevention, treatment and diagnosis of
CC   diseases associated with altered CNAP expression such as, cell
CC   proliferative disorders (e.g. arteriosclerosis, cirrhosis, leukaemia,
CC   lymphoma and cancer of the breast, prostate, lung and brain), autoimmune/
CC   inflammatory disorders (e.g. asthma, anaemia, diabetes mellitus, multiple
CC   sclerosis and psoriasis), neurological disorders (e.g. epilepsy,
CC   Alzheimer's/Parkinson's disease and strokes), vision disorders (e.g.
CC   conjunctivitis, glaucoma, cataracts and retinitis pigmentosa),
CC   reproductive disorders (e.g. infertility, uterine fibroids, ectopic
CC   pregnancies and impotence) and smooth muscle disorders (e.g. angina,
CC   anaphylactic shock, Kearns-Sayre syndrome and hypertension). The
CC   coding sequence can be used for gene therapy.
XX
SQ   Sequence   690 AA;

  Query Match             100.0%;  Score 3593;  DB 21;  Length 690;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  690;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660

Qy        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
              |||||||||||||||||||||||||||||
Db        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
```

RESULT 2
AAY51607
ID AAY51607 standard; Protein; 690 AA.
XX
AC AAY51607;
XX
DT 26-MAY-2000 (first entry)
XX
DE Human soluble guanylylcyclase alpha1 protein.
XX
KW Human; guanylylcyclase alpha1; hsGCalpha1; hsGCbeta1; soluble;
KW guanylylcyclase beta1; antiarteriosclerotic; vasotropic; hypotensive;
KW gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;
KW peripheral arterial occlusive disease; arterial hypertension.
XX
OS Homo sapiens.
XX
PN DE19837015-A1.
XX
PD 24-FEB-2000.
XX
PF 14-AUG-1998; 98DE-1037015.
XX
PR 14-AUG-1998; 98DE-1037015.
XX
PA (VASO-) VASOPHARM BIOTECH GMBH & CO KG.
XX
PI Schmidt H, Zabel U, Poller W;
XX
DR WPI; 2000-184044/17.
DR N-PSDB; AAZ88938.
XX
PT New human soluble guanylate cyclase alpha1/beta1 and the nucleic acid
PT encoding the subunits, useful for producing diagnostic antibodies, and
PT for somatic gene therapy of arteriosclerosis -
XX
PS Example 1; Page 13-15; 44pp; German.
XX
CC This invention describes novel purified human soluble guanylate cyclase
CC alpha1/beta1 (hsGCalpha1/beta1). The products of the invention have
CC antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid
CC sequences encoding the alpha1 and/or beta1 subunit are useful for somatic
CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),
CC peripheral arterial occlusive disease and arterial hypertension.
CC Antibodies to hsGCalpha1/beta1 can be used for diagnosis of aberrant
CC hsGCalpha1/beta1 expression in human tissues. This sequence represents
CC the human soluble guanylylcyclase alpha1 subunit described in the method
CC of the invention.
XX
SQ Sequence 690 AA;

Query Match 100.0%; Score 3593; DB 21; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660

Qy        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
              |||||||||||||||||||||||||||||
Db        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
```

RESULT 3
ABG61940
ID ABG61940 standard; Protein; 690 AA.
XX
AC ABG61940;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #141.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US32045.
XX
PR 13-OCT-2000; 2000US-0687576.
PR 08-DEC-2000; 2000US-0733288.
PR 08-DEC-2000; 2000US-0733742.
PR 24-JAN-2001; 2001US-263957P.
PR 16-MAR-2001; 2001US-276791P.
PR 16-MAR-2001; 2001US-276888P.
PR 06-APR-2001; 2001US-281922P.
PR 24-APR-2001; 2001US-286214P.
PR 30-APR-2001; 2001US-0847046.
PR 04-MAY-2001; 2001US-288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
DR WPI; 2002-471335/50.
DR N-PSDB; ABK92259.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes

PT are expressed in a prostate tissue -
XX
PS Claim 27; Page 421-422; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate
CC cancer-associated transcript in a cell from a patient. The method
CC comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridise to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
SQ Sequence 690 AA;

Query Match 100.0%; Score 3593; DB 23; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
Db 1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

Qy 61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
Db 61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy 121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
Db 121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180

Qy 181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
Db 181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240

Qy 241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
Db 241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300

Qy 301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
Db 301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360

Qy 361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
Db 361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy 421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
Db 421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480

Qy 481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
Db 481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540

Qy 541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
Db 541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy 601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
Db 601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660

Qy 661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
Db 661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690

RESULT 4
ABG61878
ID ABG61878 standard; Protein; 717 AA.
XX
AC ABG61878;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #79.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US32045.
XX
PR 13-OCT-2000; 2000US-0687576.
PR 08-DEC-2000; 2000US-0733288.
PR 08-DEC-2000; 2000US-0733742.
PR 24-JAN-2001; 2001US-263957P.
PR 16-MAR-2001; 2001US-276791P.
PR 16-MAR-2001; 2001US-276888P.
PR 06-APR-2001; 2001US-281922P.
PR 24-APR-2001; 2001US-286214P.
PR 30-APR-2001; 2001US-0847046.
PR 04-MAY-2001; 2001US-288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
DR WPI; 2002-471335/50.
DR N-PSDB; ABK92193.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue -
XX
PS Claim 27; Page 363; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate
CC cancer-associated transcript in a cell from a patient. The method
CC comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridise to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
SQ Sequence 717 AA;

Query Match 88.3%; Score 3171.5; DB 23; Length 717;
Best Local Similarity 90.7%; Pred. No. 1.1e-294;
Matches 625; Conservative 10; Mismatches 37; Indels 17; Gaps 4;

Qy 1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

```
Db          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQE-- 178
              |||    |||: |  | :   |    : :  || :   || |         ||  :  |
Db        121 QAVQQS-PVELSKNLLVKRFLKYVTRKMKTSLGWLEAPLKIF--------KQLQYPSETE 171

Qy        179 -----AGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEV 233
                  : |:|:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 QPLPRSRKKGQLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEV 231

Qy        234 SLMPPCFHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFD 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 SLMPPCFHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFD 291

Qy        294 KDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRW 353
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        292 KDMTILQFGNGIRRLMNRRDFQGKPNF-EYFEILTPKINQTFSGIMTMLNMQFVVRVRRW 350

Qy        354 DNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVV 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 DNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVV 410

Qy        414 LIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQ 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 LIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQ 470

Qy        474 AKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYC 533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIAMPIV 530

Qy        534 VAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPR 593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 WLGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPR 590

Qy        594 YCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGI 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 YCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGI 650

Qy        654 CHFLDAYQQGTNSKPCFQKKDVEDGNANF 682
              |||||||||||||||||||||||| :  |
Db        651 CHFLDAYQQGTNSKPCFQKKDVEDASQFF 679


RESULT 5
ABB58042
ID   ABB58042 standard; Protein; 676 AA.
XX
AC   ABB58042;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 918.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02145.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 918; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   676 AA;

  Query Match             27.6%;  Score 991.5;  DB 22;  Length 676;
  Best Local Similarity   37.8%;  Pred. No. 1.3e-85;
  Matches  247;  Conservative  117;  Mismatches  209;  Indels   81;  Gaps   20;

Qy         58 SLPQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAK--------HKIK---ESR 106
              :|   : |   : |  |  :|   |  |  | || |:   :||       ||:|   ::
Db         27 ALEDEELSDDALTLTHLQMAIQLLTAPSNEDLNTAVTSLVAKYRQNWPNIHKLKLDPQTF 86

Qy        107 KSLEREDFEKTIAE------QAVAAGVPVEVIKESLGEEVFKICYEE-DENILGVVGGTL 159
              ||    |:   | |      :| |: : |      ||||:   |     |     :|  |
Db         87 KSCANYDYLADIQELLLKMDEASASEILV-----LLGEELITCCCTGIIERAFRCLGTDL 141

Qy        160 KDFLNSFS---TLLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILP 216
              ::|| |      :||     ||      : |   :|  : :    : :   :   : :|
Db        142 QEFLGSLDGVYDVLK----LQEED----VTDTGFVCAGEGE----LIFTSERPVIAWLLL 189

Qy        217 GIIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQPYLLYSVH-----MKSTKP-SLSPSK 270
              | :||   :||: :| : : |      |   :    ||   |      |   :| |:| :
Db        190 GSLKALTRMLYKVDVNIKIEP--VEGDARRY---RYLFSLVKDNSQTMLMGRPTSVSKTI 244

Qy        271 PQ----------SSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
              |:          | | : :| ||| ||:||: :: : ::| | |  :|        ||
Db        245 PETVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLY-------KPYM 297

Qy        321 EE-------YFEILTPK-INQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIY 372
              :       ||:   || :   |  |:      |:: :     :|   :  :::||||::
Db        298 ADFGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIGLEIKGQMVH 357

Qy        373 IVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKL 432
                ||:::||:||| :| |:  |  ||::||||:|:| |:|:|:|||||||||::|: |:
Db        358 CPESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQARAQDGLRRRMDKI 417

Qy        433 KATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTA 492
              | ::|:|: |: :|:|| | ||  ||| |:|::|| |   : || : :||:|||||||:
Db        418 KNSIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYPDVTILFSDIVGFTS 477

Qy        493 ICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALM 552
              |||: :|  ||:||  ||  ||: |   |||||||||||||| |||: |   | ::| |
Db        478 ICSRATPFMVISMLEGLYKDFDEFCDFFDVYKVETIGDAYCVASGLHRASIYDAHKVAWM 537

Qy        553 ALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
              ||||:    : ::  || |||||||:|:| ||||| ||||||||::||:|||||| |
Db        538 ALKMIDACSKHITHDGEQIKMRIGLHTGTVLAGVVGRKMPRYCLFGHSVTIANKFESGSE 597

Qy        613 PRKINVSPTTYRLLKDCPGFVF--TPRSREELPPNFPSEIPG---ICHFLDAYQ 661
              ||||||||   |   || |  ||  ||  ||:  ||    |:||::::
Db        598 ALKINVSPTTKDWLTKHEGFEFELQPRDPSFLPKEFPN--PGGTETCYFLESFR 649
```

```
RESULT 6
AAY51608
ID   AAY51608 standard; Protein; 619 AA.
XX
AC   AAY51608;
XX
DT   26-MAY-2000  (first entry)
XX
DE   Human soluble guanylylcyclase beta1 protein.
XX
KW   Human; guanylylcyclase alpha1; hsGCalpha1; hsGCbeta1; soluble;
KW   guanylylcyclase beta1; antiarteriosclerotic; vasotropic; hypotensive;
KW   gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;
KW   peripheral arterial occlusive disease; arterial hypertension.
XX
OS   Homo sapiens.
XX
PN   DE19837015-A1.
XX
PD   24-FEB-2000.
XX
PF   14-AUG-1998;   98DE-1037015.
XX
PR   14-AUG-1998;   98DE-1037015.
XX
PA   (VASO-) VASOPHARM BIOTECH GMBH & CO KG.
XX
PI   Schmidt H,  Zabel U,  Poller W;
XX
DR   WPI; 2000-184044/17.
DR   N-PSDB; AAZ88939.
XX
PT   New human soluble guanylate cyclase alpha1/beta1 and the nucleic acid
PT   encoding the subunits, useful for producing diagnostic antibodies, and
PT   for somatic gene therapy of arteriosclerosis -
XX
PS   Example 1; Page 16-17; 44pp; German.
XX
CC   This invention describes novel purified human soluble guanylate cyclase
CC   alpha1/beta1 (hsGCalpha1/beta1). The products of the invention have
CC   antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid
CC   sequences encoding the alpha1 and/or beta1 subunit are useful for somatic
CC   gene therapy of arteriosclerosis and restenosis, ischemia (infarct),
CC   peripheral arterial occlusive disease and arterial hypertension.
CC   Antibodies to hsGCalpha1/beta1 can be used for diagnosis of aberrant
CC   hsGCalpha1/beta1 expression in human tissues. This sequence represents
CC   the human soluble guanylylcyclase beta1 subunit described in the method
CC   of the invention.
XX
SQ   Sequence   619 AA;

  Query Match             22.0%;  Score 789.5;  DB 21;  Length 619;
  Best Local Similarity   35.1%;  Pred. No. 2.7e-66;
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153
              :: |  : | |  :: :     |||   |      |:: :  |:  | | |   : ||
Db         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88

Qy        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212
              |:|  :::|| :   |     |   |     :   |  | | |    | ::|:  :
Db         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143

Qy        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265
               |: ||||  |  :: ||::: ::      | |             : :  | :|
Db        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192

Qy        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
                  :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250

Qy        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370
                 | :: | |: :| ||:: :|  ||:| :     |:|             : |||||
Db        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310

Qy        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430
              ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370

Qy        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490
               |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430

Qy        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543
               | ||: :     :::: :|| ||||||          ||||||:|| |    || :
Db        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490

Qy        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603
               ||  |  :|| |||:: :|    || ::: ||:|:| |  ||:| :||||||||| | |
Db        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549

Qy        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652
               :: |:     ||||  ||| |           | |   | |  |  |
Db        550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588


RESULT 7
ABB58149
ID   ABB58149 standard; Protein; 787 AA.
XX
AC   ABB58149;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 1239.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02252.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 1239; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
```

```
SQ   Sequence   787 AA;

  Query Match             20.7%;  Score 743;  DB 22;  Length 787;
  Best Local Similarity   28.8%;  Pred. No. 1.1e-61;
  Matches  198;  Conservative  110;  Mismatches  221;  Indels  158;  Gaps   15;

Qy         89 LNVALQRTLAKH-------KIKESR------KSLEREDFEKTIAEQAVAAGV-----PVE 130
              :| ||:  : ||       |||:        : | |: ::  |    : | |     | :
Db          5 VNYALELLVLKHFGEEIWEKIKKKAMVSMEGQFLVRQIYDDEITYNLIGAAVEILNIPAD 64

Qy        131 VIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDAS 189
               | |  |:  |: | :   : || |:| | :||| :   |     |          :   |
Db         65 DILELFGKTFFEFCQDSGYDKILQVLGATPRDFLQNLDAL-----HDHLGTLYPGMRAPS 119

Qy        190 ILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLM------------ 236
               | :|: : | ::|:  :     |: ||:|| |  |:  |||: ::
Db        120 FRCTEKDGELL-LHYYSERPGLEHIVIGIVKAVASKLHGVEVEIDIVKRKGEPIDEAEKE 178

Qy        237 ------------------------------------------------------------ 236

Db        179 RAIARENQQLLEDAVATTTTGSATVVLAPSSDAERNNNHNGSNGSNNNGMANNGNTVNVN 238

Qy        237 ----------------------------------PPCFH------NDCSEFVNQPYLLY 255
                                                 |       :|   : :  :|:
Db        239 NNNDGQQIASETDPSIALSTCPIAQDSFDCDGDKEQKCLRLLKNKSDDIERYDHVQFLIR 298

Qy        256 SVHM--KSTKPSLSPSKP-------QSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIR 306
              :::  ||   :     |        :: |: | : ||| |||| |||: | |:| |  :
Db        299 EINVAAKSQVDAKKDEVPDDMEFLCEAPLISPAT-FCKVFPFHLMFDRQMKIVQAGKAVS 357

Qy        307 RLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDL 366
              |:: | :   :  | | : | :   |  |:: :|  :|:: |:   | :   | : |
Db        358 RVI-PRVAEENCSLIEVVEAIRPHLQLNFENILSHINTIYVLQTRQGAMSSRHEQRFLRL 416

Qy        367 KGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLK 426
              ||||:|| |:  |||   | |  |:| | :|||:||:|:|:| ||:||: |:  |:  |
Db        417 KGQMMYIPETDRILFQCYPSVMNLDDLTKKGLYISDVPLHDAARDLVLLSEKFEAEYKLT 476

Qy        427 KRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSD 486
              | |  |   |:|  : || ||:||  || |: |  || :|   : |  |:: :||::||
Db        477 KNLEMLTDKLQQTFRDLESEKQKTDRLLYSVLPKSVANELRHQRPVPPKRYDSVTLMFSG 536

Qy        487 IVGFTAICSQCS----PLQVITMLNALYTRFD---QQCGELDVYKVETIGDAYCVAGGLH 539
              ||||   |:  :     :::: ||| ||| ||        |:||||||:|| |    ||
Db        537 IVGFGQYCAANTDPDGAMKIVKMLNELYTVFDALTDSKRNLNVYKVETVGDKYMAVSGLP 596

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
               : ||  :| :|| ||:::  |      |::: ||:||| |  ||:| ::||||||||
Db        597 DHCEDHAKCMARVALDMMDMAKNV-KMGSNPVQITIGIHSGEVVTGVIGNRVPRYCLFGN 655

Qy        600 NVTLANKFESCSVPRKINVSPTTYRLL 626
              | | :: |:  || :||||  |||||
Db        656 TVNLTSRTETTGVPGRINVSEETYRLL 682


RESULT 8
ABB60360
ID   ABB60360 standard; Protein; 690 AA.
XX
AC   ABB60360;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 7872.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL04463.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 7872; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   690 AA;

  Query Match             17.9%;  Score 642.5;  DB 22;  Length 690;
  Best Local Similarity   30.6%;  Pred. No. 4e-52;
  Matches  177;  Conservative  107;  Mismatches  200;  Indels   95;  Gaps   18;

Qy        149 ENILGVVGGTLKDFLNSFSTL---LKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYF 205
              : :| |:|  ::||||    |   || |          |:  | :| ::    | ::|
Db         83 DRVLSVLGRHMRDFLNGLDNLHEYLKFS--------YPRMRAPSFICENETKQGLTLHYR 134

Qy        206 FPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQPYLLYSVHM------ 259
                :|       | |:  |   |  |: :              | | :  |  :||:
Db        135 SKRRGFVYYTMGQIREVARYFYHKEMHI-----------ELVREEILFDTVHVTFQLTF 182

Qy        260 ---KSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQG 316
                 |  ||: :: :  | |   :  : |||  :|  || :   || :  ::   :  |
Db        183 DNRAFTLASLAMTREEKHLPISAHVLFEIFPFCMVFGADMVVRSIGNSLMVIL--PELLG 240

Qy        317 KPNFEEYFEILTPKINQTFSGIMTMLNMQF-VVRV----RRWD-----------NSVKK 359
              |     :|::: | |   |  |:   |  | :| |     |:|            :  :|
Db        241 K-KITAWFDLVRPLIAFKFQTILNRTNNIFELVTVDPVTERFDVQNEDLLQHEDGSEPEK 299

Qy        360 SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQA 419
              | |   |||||:|:     |:|||:| :  |       |||::|: :|:  ||::| | |
Db        300 SLR---LKGQMVYMENWRMIMFLGTPVMPDLTSLITTGLYINDLSMHDFSRDLMLAGTQQ 356

Qy        420 RAQDGLKKRLG---KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQ--VVQA 474
                :  ||  |    :    ||:: : |:|| ::| :||  : | :|| :| :|:  :
Db        357 SVE--LKLALDQEQQKSKKLEESMRLLDEEMRRTDELLYQMIPKQVADRLRRGENPIDTC 414

Qy        475 KKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCV 534
              : | :|::|||||| || |||: :|::|::||||:|: ||:      |||||||||| |
Db        415 EMFDSVSILFSDIVTFTEICSRITPMEVVSMLNAMYSIFDKLTERNSVYKVETIGDAYMV 474

Qy        535 AGGLHKESDTHAVQIALMALKMMELSDEVMSPH-GEPIKMRIGLHSGSVFAGVVGVKMPR 593
               |    :   || ::  |||  |::   ::   |  |: :::|:|:|:|||:| ||:||||||
Db        475 VAGAPDKDANHAERVCDMALDMVDAITDLKDPSTGQHLRIRVGVHSGAVVAGIVGLKMPR 534

Qy        594 YCLFGNNVTLANKFESCSVPRKINVSPTT-------YRLL---------KDCPGFVFTPR 637
              |||||: |  |:: || |:  |:::| :|       |:::         |   |  :
Db        535 YCLFGDTVNTASRMESTSIAMKVHISESTKVLIGPNYKIIERGEIDVKGKGTMGTYWLEE 594
```

```
Qy        638 SREELP----------PNFPSEIPGICHFLDAYQQGTNS 666
              ||          |  ||      | |    ||  |
Db        595 RENRLPLQLTAALQVHPLSPS------HLLPRDLQGCGS 627


RESULT 9
ABB59684
ID   ABB59684 standard; Protein; 1076 AA.
XX
AC   ABB59684;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 5844.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL03787.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 5844; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   1076 AA;

  Query Match             12.9%;  Score 462;  DB 22;  Length 1076;
  Best Local Similarity   43.6%;  Pred. No. 1.7e-34;
  Matches  109;  Conservative   38;  Mismatches   91;  Indels   12;  Gaps    5;

Qy        404 PIHNALRD---VVLIGEQARAQDGLKKRL----GKLKATLEQAHQALEEEKKKTVDLLCS 456
              |  : :|     :: |      | |  |:      |:: :|:  : |  ||::| :||
Db        785 PTFSTIRSNIRTIMKGFCENLMDDLLNRMEQYANNLESLVEEKTRQLSLEKQRTEELLYQ 844

Qy        457 IFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQ 516
              | |  |||||  | :|: ::||:||: |||||||| :|:: ||: |:  || ||: ||:
Db        845 ILPRPVAQQLMAGDLVEPEEFSSVTIYFSDIVGFTELCARSSPMDVVNFLNDLYSTFDRI 904

Qy        517 CGELDVYKVETIGDAYCVAGGL-HKESDTHAVQIALMALKMMELSDEVMSPHGEP---IK 572
              |  |||||||||||| |  ||     | || :|||||| ::          | :|   |:
Db        905 IGFYDVYKVETIGDAYLVVSGLPEPNGDKHAREIALMALDILRAVSSFNLRH-KPEYKIQ 963

Qy        573 MRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGF 632
              :|||:|||||  |||||  ||| |||||: |  |:: ||   | ||:||  |  :|     |
Db        964 IRIGMHSGSVCAGVVGKKMPHYCLFGDTVNTASRMESTGQPGKIHVSSATKAILDKFGTF 1023

Qy        633 VFTPRSREEL 642
                  |   ||
Db       1024 QMEQRGDVEL 1033


RESULT 10
AAB83940
ID   AAB83940 standard; Protein; 1115 AA.
XX
AC   AAB83940;
XX
DT   23-JUL-2001  (first entry)
XX
DE   Amino acid sequence of a guanylyl cyclase-2 chemoreceptor (HG-gcy-2).
XX
KW   Soybean cyst nematode; guanylyl cyclase-1 chemoreceptor; HG-gcy-1;
KW   parasitic nematode; chemotaxis; HG-gcy-2.
XX
OS   Heterodera glycines.
XX
PN   WO200134791-A2.
XX
PD   17-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US30295.
XX
PR   05-NOV-1999;   99US-0435376.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Davis EL,  Yan Y;
XX
DR   WPI; 2001-335924/35.
DR   N-PSDB; AAF89878.
XX
PT   Novel isolated polynucleotide encoding nematode guanylyl cyclase
PT   chemoreceptor, useful for controlling plant feeding nematodes  -
XX
PS   Example 2; Page 55-59; 69pp; English.
XX
CC   The present sequence represents soybean cyst nematode guanylyl cyclase-2
CC   chemoreceptor, designated HG-gcy-2. HG-gcg-1 polynucleotides and
CC   polypeptides are useful as active agents for controlling plant
CC   feeding nematodes. They can be used to screen for compounds that have
CC   the ability to disrupt parasitic nematode feeding or chemotaxis.
XX
SQ   Sequence   1115 AA;

  Query Match             12.8%;  Score 461;  DB 22;  Length 1115;
  Best Local Similarity   40.6%;  Pred. No. 2.2e-34;
  Matches  114;  Conservative   38;  Mismatches  101;  Indels   28;  Gaps    6;

Qy        376 SSAILFLGSPCVD-------RLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKR 428
              | |:| |   | |       ::|  |            ||   :  |      | :
Db        812 SPALLHLIKDCWDESPAERPKMETVT------------ALLQSMNTGRSTNLMDHVFNM 858

Qy        429 L----GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLF 484
              |    | |:  :|:  : | ||||||  ||  : | :|| :|  || |: : |  ||: |
Db        859 LEVYAGSLEEEVEERTKELVEEKKKTDILLYRMLPKQVADKLKLGQSVEPETFDCVTVFF 918

Qy        485 SDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAY-CVAGGLHKESD 543
              ||:| || | |:|||||||: :|| |||  |    | ||||||||| | ||:|  |:
Db        919 SDVVSFTTIASKCSPLQVVNLLNNLYTLLDSIIAEFDVYKVETIGDGYLCVSGLPHRNGH 978

Qy        544 THAVQIALMALKMMELSDEVMSPH--GEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNV 601
              ||  || |:|  |        ||   | :::|||:|:||  |||||: |||||||: :
Db        979 EHAQHIAKMSLAFMRNLGSFTIPHLPIERLRLRIGIHTGSTVAGVVGLSMPRYCLFGDTI 1038

Qy        602 TLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
                | : || | | :|::| ||   | :  |  |  ::| |:
```

```
Db       1039 NTAARLESSSKPMRIHISTTTNHFLVNVLG-GFVTQARGEI 1078


RESULT 11
AAR38863
ID   AAR38863 standard; Protein; 1025 AA.
XX
AC   AAR38863;
XX
DT   08-FEB-1994  (first entry)
XX
DE   GC-B.
XX
KW   Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
KW   mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
KW   binding domain; antibody; diarrhoea; ligand; antagonist.
XX
OS   Rattus rattus.
XX
PN   US5237051-A.
XX
PD   17-AUG-1993.
XX
PF   06-DEC-1990;   90US-0623033.
XX
PR   06-DEC-1990;   90US-0623033.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Garbers DL,  Schulz S;
XX
DR   WPI; 1993-272183/34.
XX
PT   New purified enterotoxin receptor protein - used to develop
PT   prods. for treating abnormal conditions caused by bacterially
PT   released enterotoxin, partic. diarrhoea
XX
PS   Disclosure; Fig 3; 26pp; English.
XX
CC   The sequences given in AAR38862-63 represent the guanylyl cyclases,
CC   GC-A and GC-B, which binds heat stable enterotoxin.  These proteins
CC   are enterotoxin receptors which may be used as a therapeutic to control
CC   intestinal fluid permeation as well as abnormal conditions caused
CC   by bacterially released enterotoxin.  The binding domain of the
CC   proteins, or antibodies to the proteins, can be used to eliminate
CC   diarrhoea.  The proteins may be used to isolate ligands and to screen
CC   for antagonists of toxin binding.  This sequence is given as it is
CC   represented in the specification.
XX
SQ   Sequence   1025 AA;

  Query Match             12.8%;  Score 460;  DB 14;  Length 1025;
  Best Local Similarity   41.9%;  Pred. No. 2.4e-34;
  Matches  113;  Conservative   39;  Mismatches   76;  Indels   42;  Gaps    6;

Qy       411 DVVLIGEQARAQD-----------GLKKRLGK--------------------LKATLEQA 439
             ::||: |:  |||           |  :|  |                     |:  :|:
Db       738 ELVLLMERCWAQDPTERPDFGQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEER 797

Qy       440 HQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSP 499
             ||  |||:|   ||  | |  ||:|| :|: |||: | :||: |||||||||: :: :|
Db       798 TQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAESTP 857

Qy       500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL-HKESDTHAVQIALMALKMME 558
             :||:|:|| ||| ||      ||||||||||| |  ||  :    || :|| ||| ::
Db       858 MQVVTLLNDLYTCFDAIIDNFDVYKVETIGDAYMVVSGLPGRNGQRHAPEIARMALALL- 916

Qy       559 LSDEVMS------PHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
               | |       || : :::|||:|:| | |||||:||||||||: |  |:: ||
Db       917 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGQ 973

Qy       613 PRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
               ||:|| ||   | :   |    |   |:
```

```
Db       974 ALKIHVSSTTKDALDELGCFQLELRGDVEM 1003


RESULT 12
AAR10399
ID   AAR10399 standard; Protein; 1047 AA.
XX
AC   AAR10399;
XX
DT   09-APR-1991  (first entry)
XX
DE   Human Natriuretic Peptide Receptor B.
XX
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /label= signal sequence
FT   Protein         12
FT                   /label= mature NPBR
FT   Domain          23..455
FT                   /label= extracellular domain
FT                   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT                   /label= transmembrane domain
FT   Domain          479..1047
FT                   /label= cytoplasmic domain
FT                   /note= "GC and protien kinase activity"
FT   Modified-site   24..26
FT                   /label= N-glycos_site
FT   Modified-site   35..37
FT                   /label= N-glycos_site
FT   Modified-site   161..163
FT                   /label= N-glycos_site
FT   Modified-site   195..197
FT                   /label= N-glycos_site
FT   Modified-site   244..246
FT                   /label= N-glycos_site
FT   Modified-site   277..279
FT                   /label= N-glycos_site
FT   Modified-site   349..351
FT                   /label= N-glycos_site
FT   Modified-site   600..602
FT                   /label= N-glycos_site
XX
PN   WO9100292-A.
XX
PD   10-JAN-1991.
XX
PF   22-JUN-1990;   90WO-US03586.
XX
PR   23-JUN-1989;   89US-0370673.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chang M,  Goeddel D,  Lowe D;
XX
DR   WPI; 1991-036711/05.
DR   N-PSDB; AAQ10324.
XX
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
XX
PS   Claim 3; Fig 1; 49pp; English.
XX
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr-
```

```
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.  An analogue, NPRBDKC, comprising only AAs 1-433 of
CC   the mature protein, is also claimed, as well as variant described
CC   in AAR10867.
XX
SQ   Sequence   1047 AA;

  Query Match             12.8%;  Score 460;  DB 12;  Length 1047;
  Best Local Similarity   41.9%;  Pred. No. 2.5e-34;
  Matches  113;  Conservative   39;  Mismatches   76;  Indels   42;  Gaps    6;

Qy        411 DVVLIGEQARAQD-----------GLKKRLGK-------------------LKATLEQA 439
              ::||: |:  |||           |  :|  |                    |:  :|:
Db        760 ELVLLMERCWAQDPAERPDFGQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEER 819

Qy        440 HQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSP 499
               ||  |||:|   ||  | |  ||:|| :|: ||| | :||: ||||||||| :: :|
Db        820 TQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAESTP 879

Qy        500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL-HKESDTHAVQIALMALKMME 558
              :||:|:|| ||| ||      |||||||||||| |  ||  :    || :|| ||| ::
Db        880 MQVVTLLNDLYTCFDAIIDNFDVYKVETIGDAYMVVSGLPGRNGQRHAPEIARMALALL- 938

Qy        559 LSDEVMS------PHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
                | | |       || : :::|||:|:| | |||||:|||||||||: |  |:: ||
Db        939 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGQ 995

Qy        613 PRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
               ||:|| ||   | :   |    |   |:
Db        996 ALKIHVSSTTKDALDELGCFQLELRGDVEM 1025


RESULT 13
AAR10867
ID   AAR10867 standard; Protein; 1047 AA.
XX
AC   AAR10867;
XX
DT   09-APR-1991  (first entry)
XX
DE   NPRB(Pro655, Glu656, Leu663, Phe664, Ala682).
XX
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /label= signal sequence
FT   Protein         12
FT                   /label= mature NPBR
FT   Domain          23..455
FT                   /label= extracellular domain
FT                   /note= "binds natriuretic peptides A,B and C"
FT   Domain          456..456
FT                   /label= transmembrane domain
FT   Domain          479..1047
FT                   /label= cytoplasmic domain
FT                   /note= "GC and protien kinase activity"
FT   Modified-site   24..26
FT                   /label= N-glycos_site
FT   Modified-site   35..37
FT                   /label= N-glycos_site
FT   Modified-site   161..163
FT                   /label= N-glycos_site
FT   Modified-site   195..197
FT                   /label= N-glycos_site
FT   Modified-site   244..246
FT                   /label= N-glycos_site
FT   Modified-site   277..279
FT                   /label= N-glycos_site
FT   Modified-site   349..351
FT                   /label= N-glycos_site
FT   Modified-site   600..602
FT                   /label= N-glycos_site
XX
PN   WO9100292-A.
XX
PD   10-JAN-1991.
XX
PF   22-JUN-1990;   90WO-US03586.
XX
PR   23-JUN-1989;   89US-0370673.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chang M,  Goeddel D,  Lowe D;
XX
DR   WPI; 1991-036711/05.
XX
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
XX
PS   Claim 27; Disclsoure; 49pp; English.
XX
CC   The sequence is that of a variant of natriuretic peptide receptor B
CC   (NPRB).
CC   Editor's note: the AAs replaced in the sequence do not correspond
CC   to those of NPRB given in the specification (see p.27, l.23-27;
CC   figs. 1 and 2), but instead to AAs in the sequence provided for
CC   NPRA (fig. 2).  The origin of the AAs to be substituted is not
CC   clear; they may be taken from the sequence of rat NPRB (p.8, l.18;
CC   p.9, l.33) or rat NPRA (p.27, l.20).  The substitution indicated on
CC   p.9, l.33 , i.e. "Phe644" is thought to be an error and should read
CC   "Phe664".
CC   The mutant was produced by modifying the DNA encoding the
CC   native sequence. The protein has guanyl cyclase (GC) activity and
CC   protein kinase activity.  The DNA can be inserted into expression
CC   vectors for the prodn. of the NPRB analogue.  The protein can be
CC   used in treatment of natriuretic peptide disorders, and also to
CC   isolate peptides using affinity chromatography.  Antibodies with
CC   affinity for NPRB can also be prepd.
CC   See also AAR10399.
XX
SQ   Sequence   1047 AA;

  Query Match             12.8%;  Score 460;  DB 12;  Length 1047;
  Best Local Similarity   41.9%;  Pred. No. 2.5e-34;
  Matches  113;  Conservative   39;  Mismatches   76;  Indels   42;  Gaps    6;

Qy        411 DVVLIGEQARAQD-----------GLKKRLGK-------------------LKATLEQA 439
              ::||: |:  |||           |  :|  |                    |:  :|:
Db        760 ELVLLMERCWAQDPAERPDFGQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEER 819

Qy        440 HQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSP 499
               ||  |||:|   ||  | |  ||:|| :|: ||| | :||: ||||||||| :: :|
Db        820 TQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAESTP 879

Qy        500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL-HKESDTHAVQIALMALKMME 558
              :||:|:|| ||| ||      |||||||||||| |  ||  :    || :|| ||| ::
Db        880 MQVVTLLNDLYTCFDAIIDNFDVYKVETIGDAYMVVSGLPGRNGQRHAPEIARMALALL- 938

Qy        559 LSDEVMS------PHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
                | | |       || : :::|||:|:| | |||||:|||||||||: |  |:: ||
Db        939 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGQ 995

Qy        613 PRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
               ||:|| ||   | :   |    |   |:
Db        996 ALKIHVSSTTKDALDELGCFQLELRGDVEM 1025


RESULT 14
AAR38862
```

```
ID   AAR38862 standard; Protein; 1029 AA.
XX
AC   AAR38862;
XX
DT   08-FEB-1994  (first entry)
XX
DE   GC-A.
XX
KW   Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
KW   mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
KW   binding domain; antibody; diarrhoea; ligand; antagonist.
XX
OS   Rattus rattus.
XX
PN   US5237051-A.
XX
PD   17-AUG-1993.
XX
PF   06-DEC-1990;   90US-0623033.
XX
PR   06-DEC-1990;   90US-0623033.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Garbers DL,  Schulz S;
XX
DR   WPI; 1993-272183/34.
XX
PT   New purified enterotoxin receptor protein - used to develop
PT   prods. for treating abnormal conditions caused by bacterially
PT   released enterotoxin, partic. diarrhoea
XX
PS   Disclosure; Fig 3; 26pp; English.
XX
CC   The sequences given in AAR38862-63 represent the guanylyl cyclases,
CC   GC-A and GC-B, which binds heat stable enterotoxin.  These proteins
CC   are enterotoxin receptors which may be used as a therapeutic to control
CC   intestinal fluid permeation as well as abnormal conditions caused
CC   by bacterially released enterotoxin.  The binding domain of the
CC   proteins, or antibodies to the proteins, can be used to eliminate
CC   diarrhoea.  The proteins may be used to isolate ligands and to screen
CC   for antagonists of toxin binding.
XX
SQ   Sequence   1029 AA;

  Query Match             12.7%;  Score 457.5;  DB 14;  Length 1029;
  Best Local Similarity   44.9%;  Pred. No. 4.2e-34;
  Matches  102;  Conservative   40;  Mismatches   78;  Indels    7;  Gaps    3;

Qy        423 DGLKKRL----GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFS 478
              | |  |:      |:  :|:  ||  |||:|   ||  | |  ||:|| :|: |||: |
Db        782 DNLLSRMEQYANNLEELVEERTQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFD 841

Qy        479 NVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL 538
              :||: |||||||||: :: :|:||:|:|| ||| ||      ||||||||||| |  ||
Db        842 SVTIYFSDIVGFTALSAESTPMQVVTLLNDLYTCFDAVIDNFDVYKVETIGDAYMVVSGL 901

Qy        539 H-KESDTHAVQIALMALKMMEL--SDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYC 595
               :    || ::| ||| :::   |  :     | :::|||:|:| | |||||:||||||
Db        902 PVRNGQLHAREVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYC 961

Qy        596 LFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              |||: |  |:: ||     ||::|  |  :|::  ||    |   |:
Db        962 LFGDTVNTASRMESNGEALKIHLSSETKAVLEEFDGFELELRGDVEM 1008


RESULT 15
AAU11280
ID   AAU11280 standard; Protein; 1061 AA.
XX
AC   AAU11280;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human atrionatriuretic peptide receptor A (NPR1) protein.
XX
KW   Human; natriuretic peptide receptor A/guanylate cyclase A; NPR1;
KW   atrionatriuretic peptide receptor A; haplotyping; cytostatic; genotyping;
KW   haplotype pair; single nucleotide polymorphism; gene therapy;
KW   drug screening; hypertension; hypotensive.
XX
OS   Homo sapiens.
XX
PN   WO200179231-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US12300.
XX
PR   14-APR-2000; 2000US-197330P.
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Bentivegna SC,  Choi JY,  Kliem SE,  Nandabalan K;
XX
DR   WPI; 2002-066340/09.
DR   N-PSDB; AAS16995.
XX
PT   Genotyping human natriuretic peptide receptor A/guanylate cyclase gene
PT   of an individual, involves determining identity of nucleotide pair at
PT   specific polymorphic sites for two copies of the gene  -
XX
PS   Claim 27; Fig 3; 96pp; English.
XX
CC   The invention relates to single nucleotide polymorphisms in the gene
CC   encoding the human natriuretic peptide receptor A/guanylate cyclase A
CC   (atrionatriuretic peptide receptor A) or NPR1 polypeptide. A method for
CC   haplotyping the NPR1 gene in an individual comprises identifying the
CC   nucleotide at one or more polymorphic sites and determining whether one
CC   of the copies of the gene is defined by one of the NPR1 haplotypes given
CC   in the specification or whether both copies are defined by a haplotype
CC   pair. This method is useful in genotyping, whereby all possible haplotype
CC   pairs can be assigned to specific genotypes. An association between a
CC   trait and a haplotype or haplotype pair of the NPR1 gene can be
CC   identified by comparing the frequency of the haplotype or haplotype pair
CC   in a population exhibiting the trait with the frequency of the haplotype
CC   or haplotype pair in a reference population, where a higher haplotype
CC   frequency in the trait population indicates the trait is associated with
CC   the haplotype or haplotype pair. NPR1 and its corresponding DNA are used
CC   for studying the expression and function of NPR1, for use in screening
CC   for candidate drugs to treat diseases related to NPR1 activity, such as
CC   hypertension. The sequences are also useful for studying the effect of
CC   variation on the biological activity of NPR1 as well as on the binding
CC   affinity of candidate drugs targeting NPR1. This sequence represents the
CC   human NPR1 polypeptide.
XX
SQ   Sequence   1061 AA;

  Query Match             12.7%;  Score 456.5;  DB 23;  Length 1061;
  Best Local Similarity   44.9%;  Pred. No. 5.5e-34;
  Matches  102;  Conservative   40;  Mismatches   78;  Indels    7;  Gaps    3;

Qy        423 DGLKKRL----GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFS 478
              | |  |:      |:  :|:  ||  |||:|   ||  | |  ||:|| :|: |||: |
Db        814 DNLLSRMEQYANNLEELVEERTQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFD 873

Qy        479 NVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL 538
              :||: |||||||||: :: :|:||:|:|| ||| ||      ||||||||||| |  ||
Db        874 SVTIYFSDIVGFTALSAESTPMQVVTLLNDLYTCFDAVIDNFDVYKVETIGDAYMVVSGL 933

Qy        539 H-KESDTHAVQIALMALKMMEL--SDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYC 595
               :    || ::| ||| :::   |  :     | :::|||:|:| | |||||:||||||
Db        934 PVRNGRLHACEVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYC 993

Qy        596 LFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              |||: |  |:: ||     ||::|  |  :|::  ||    |   |:
```

Db 994 LFGDTVNTASRMESNGEALKIHLSSETKAVLEEFGGFELELRGDVEM 1040

Search completed: June 27, 2003, 13:03:00
Job time : 39.3713 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:07 ; Search time 19.5034 Seconds
(without alignments)
3401.080 Million cell updates/sec

Title: US-09-762-767A-2
Perfect score: 3593
Sequence: 1 MFCTKLKDLKITGECPFSLL.....QKKDVEDGNANFLGKASGID 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3232	90.0	690	1	OYRTA1	guanylate cyclase
2	3171.5	88.3	717	2	S23098	guanylate cyclase
3	3107.5	86.5	691	1	OYBO77	guanylate cyclase
4	1635	45.5	732	2	S18325	guanylate cyclase,
5	976.5	27.2	683	2	JH0810	guanylate cyclase
6	795	22.1	619	1	OYRTB1	guanylate cyclase
7	793.5	22.1	619	1	OYBO70	guanylate cyclase
8	789.5	22.0	619	1	S23097	guanylate cyclase
9	774.5	21.6	682	1	OYRTB2	guanylate cyclase
10	749.5	20.9	751	2	T24458	hypothetical prote
11	587	16.3	685	2	T19968	hypothetical prote
12	582.5	16.2	686	2	T23721	hypothetical prote
13	581	16.2	1099	2	T16822	hypothetical prote
14	551.5	15.3	699	2	T18984	hypothetical prote
15	493.5	13.7	583	2	F88642	protein C54E4.3 [i
16	463.5	12.9	1057	2	T15720	hypothetical prote
17	460	12.8	1047	1	OYHUBR	natriuretic peptid
18	460	12.8	1047	1	OYRTBR	atrial natriuretic
19	460	12.8	1047	2	I45882	guanylate cyclase
20	457.5	12.7	1057	1	OYRTR	atrial natriuretic
21	457.5	12.7	1057	2	I55319	guanylyl cyclase A
22	456.5	12.7	1061	1	OYHUAR	natriuretic peptid
23	454.5	12.6	1057	2	I57963	natriuretic peptid
24	452	12.6	1100	2	T42260	guanylate cyclase
25	449.5	12.5	1057	1	OYMSAR	atrial natriuretic
26	448	12.5	1108	2	A55915	guanylate cyclase
27	442	12.3	1108	2	I59385	guanylate cyclase
28	440	12.2	1103	2	JC5581	guanylate cyclase
29	438	12.2	1108	2	B55915	guanylate cyclase
30	438	12.2	1110	1	S55279	guanylate cyclase
31	436	12.1	632	2	T31666	natriuretic peptid
32	435	12.1	1102	2	JH0717	guanylate cyclase
33	434	12.1	632	2	T31667	guanylate cyclase
34	430	12.0	1110	2	I59370	guanylate cyclase
35	428.5	11.9	1050	2	S45636	natriuretic-peptid
36	428.5	11.9	1130	2	A89130	protein F52E1.4 [i
37	428	11.9	1068	2	T42382	guanylate cyclase
38	427	11.9	1125	1	OYURCP	speract receptor p
39	423.5	11.8	1525	1	A56699	guanylate cyclase
40	420.5	11.7	1005	2	S33525	guanylate cyclase
41	418	11.6	1012	2	T24384	hypothetical prote
42	418	11.6	1170	2	T27564	hypothetical prote
43	414	11.5	1144	2	T21223	hypothetical prote
44	412.5	11.5	1122	2	T28130	hypothetical prote
45	411.5	11.5	540	2	T34187	hypothetical prote

ALIGNMENTS

RESULT 1
OYRTA1
guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - rat
N;Alternate names: guanylate cyclase, soluble, 77K chain
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A38297
R;Nakane, M.; Arai, K.; Saheki, S.; Kuno, T.; Buechler, W.; Murad, F.
J. Biol. Chem. 265, 16841-16845, 1990
A;Title: Molecular cloning and expression of cDNAs coding for soluble guanylate cycla
A;Reference number: A38297; MUID:91009100; PMID:1698769
A;Accession: A38297
A;Molecule type: mRNA
A;Residues: 1-690 <NAK>
A;Cross-references: GB:M57405; GB:M36075; NID:g204277; PIDN:AAA41206.1; PID:g204278
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;432-660/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 90.0%; Score 3232; DB 1; Length 690;
Best Local Similarity 89.1%; Pred. No. 4.6e-217;
Matches 616; Conservative 39; Mismatches 34; Indels 2; Gaps 2;

```
Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||| | |||||||||||||||||| |  || || |||||:||:| ||:  | | : | |
Db          1 MFCRKFKDLKITGECPFSLLAPGQVPTEPIEEVAGVSESCQATLPTCQEFAE-NAEGSHP 59

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||:|||||||||||| ||||||||||:|||||||||||||:| | |:|| |: |||
Db         60 QRKTSRNRVYLHTLAESIGKLIFPEFERLNLALQRTLAKHKIKENRNSSEKEDLERIIAE 119

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              :|:||||||||:|:|||||:||||||||:|||||||||||||||||||||||||||||||
Db        120 EAIAAGVPVEVLKDSLGEELFKICYEEDEHILGVVGGTLKDFLNSFSTLLKQSSHCQEAE 179

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              :|||||||||||||: |||:||||||||||:|:|||||||| :|||: ||||||||||
Db        180 RRGRLEDASILCLDKDQDFLNVYYFFPKRTTALLLPGIIKAAARILYESHVEVSLMPPCF 239

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              ::|:|||||||||||||:|||||||| |||||||||||||||||||||| |:|: |||
Db        240 RSECTEFVNQPYLLYSVHVKSTKPSLSPGKPQSSLVIPTSLFCKTFPFHFMLDRDLAILQ 299

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              |||||||:|:|||||||||||:|||||||||||||||||||||||:||||||| ||||
Db        300 LGNGIRRLVNKRDFQGKPNFEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRWDNLVKKS 359

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 419

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
```

```
                ||||||||||||||||| |||||||||||||||||| |||||||||:||||||: |
Db        420 AQDGLKKRLGKLKATLEHAHQALEEEKKKTVDLLCSIFPSEVAQQLWQGQIVQAKKFNEV 479

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        480 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHR 539

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        540 ESDTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 599

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        600 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLDAY 659

Qy        661 Q-QGTNSKPCFQKKDVEDGNANFLGKASGID 690
              | || |||| ||:|| |||||||||||||:|
Db        660 QHQGPNSKPWFQQKDAEDGNANFLGKASGVD 690


RESULT 2
S23098
guanylate cyclase (EC 4.6.1.2), soluble, 81K chain - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S23098
R;Giuili, G.; Scholl, U.; Bulle, F.; Guellaen, G.
FEBS Lett. 304, 83-88, 1992
A;Title: Molecular cloning of the cDNAs coding for the two subunits of soluble guanylyl
A;Reference number: S23097; MUID:92316204; PMID:1352257
A;Accession: S23098
A;Molecule type: mRNA
A;Residues: 1-717 <GIU>
A;Cross-references: EMBL:X66534; NID:g31683; PIDN:CAA47145.1; PID:g31684
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: heterodimer; phosphorus-oxygen lyase
F;430-658/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             88.3%;  Score 3171.5;  DB 2;  Length 717;
  Best Local Similarity   90.7%;  Pred. No. 7.9e-213;
  Matches  625;  Conservative   10;  Mismatches   37;  Indels   17;  Gaps    4;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQE-- 178
              |||    |||: |  | :   |    : :  || :  || |         ||  :  |
Db        121 QAVQQS-PVELSKNLLVKRFLKYVTRKMKTSLGWLEAPLKIF--------KQLQYPSETE 171

Qy        179 -----AGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEV 233
                  : |:|:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 QPLPRSRKKGQLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEV 231

Qy        234 SLMPPCFHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFD 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 SLMPPCFHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFD 291

Qy        294 KDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRW 353
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        292 KDMTILQFGNGIRRLMNRRDFQGKPNF-EYFEILTPKINQTFSGIMTMLNMQFVVRVRRW 350

Qy        354 DNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVV 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 DNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVV 410

Qy        414 LIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQ 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 LIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQ 470

Qy        474 AKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYC 533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIAMPIV 530

Qy        534 VAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPR 593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 WLGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPR 590

Qy        594 YCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGI 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 YCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGI 650

Qy        654 CHFLDAYQQGTNSKPCFQKKDVEDGNANF 682
              |||||||||||||||||||||||| :  |
Db        651 CHFLDAYQQGTNSKPCFQKKDVEDASQFF 679


RESULT 3
OYBO77
guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - bovine
N;Alternate names: guanylate cyclase, soluble, 77K chain
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
C;Accession: S10713; A38767
R;Koesling, D.; Harteneck, C.; Humbert, P.; Bosserhoff, A.; Frank, R.; Schultz, G.; B
FEBS Lett. 266, 128-132, 1990
A;Title: The primary structure of the larger subunit of soluble guanylyl cyclase from
A;Reference number: S10713; MUID:90306336; PMID:1973124
A;Accession: S10713
A;Molecule type: mRNA
A;Residues: 1-691 <KOE1>
A;Cross-references: EMBL:X54014; NID:g405; PIDN:CAA37960.1; PID:g406
A;Accession: A38767
A;Molecule type: protein
A;Residues: 118-133;226-232;286-293;319-330;412-417;557-571;629-637 <KOE2>
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;434-662/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             86.5%;  Score 3107.5;  DB 1;  Length 691;
  Best Local Similarity   86.1%;  Pred. No. 2.1e-208;
  Matches  595;  Conservative   41;  Mismatches   54;  Indels    1;  Gaps    1;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVP-ICQDIPEKNIQESL 59
              ||| |||||:|||:|||||||||||| |   || ||  :     | :|  :|:||    |
Db          1 MFCAKLKDLQITGDCPFSLLAPGQVPREPLGEATGSGPASTPGQPGVCPGVPDKNPPGRL 60

Qy         60 PQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIA 119
              |:|||||||||||||||||||||||||||:||||||||||||:||||||||||| :
Db         61 PRRKTSRSRVYLHTLAESICKLIFPEFERLNLALQRTLAKHKIKENRKSLEREDFEKIVV 120

Qy        120 EQAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEA 179
              :||:|||||||:||||||||:|||||||| ||||||||||||||||||||||||||||||
Db        121 DQAIAAGVPVEIIKESLGEELFKICYEEDEYILGVVGGTLKDFLNSFSTLLKQSSHCQEA 180

Qy        180 GKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPC 239
              |:|| |||||||||: | |:|||||||| |||||||||||| :|||||||| |
Db        181 EKKGRFEDASILCLDKDPDVLYVYYFFPKRITSLILPGIIKAAARILYETEVEVSSTPSR 240

Qy        240 FHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTIL 299
              || || |||:||  |||||::| :|   | || |||||| |||||||||||| |:||:||
Db        241 FHQDCREFVDQPCELYSVHIRSARPHPPPGKPVSSLVIPASLFCKTFPFHFMLDRDMSIL 300

Qy        300 QFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK 359
              | |:||||||:||| ||||:|:||||||||||:|||||||||||||:|||||||||:||
Db        301 QLGHGIRRLMSRRDVQGKPHFDEYFEILTPKISQTFSGIMTMLNMQFLVRVRRWDNSMKK 360

Qy        360 SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQA 419
              |||||||||||||:||||:|||||||||||||||||||||||||||||||||||||||||
Db        361 SSRVMDLKGQMIYMVESSSILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQA 420

Qy        420 RAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSN 479
```

```
Db        421 RAQDGLKKRLGKLKATLEQAHQALEEEKRKTVDLLCSIFPSEVARQLWQGHAVQAKRFGN 480

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539

Db        481 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDRQCGELDVYKVETIGDAYCVAGGLH 540

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599

Db        541 KESDTHAVQIALMALKMMELSHEVVSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 600

Qy        600 NVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDA 659

Db        601 NVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLEA 660

Qy        660 YQQGTNSKPCFQKKDVEDGNANFLGKASGID 690

Db        661 YQQGTTSKPWFQKKDVEEANANFLGKASGID 691


RESULT 4
S18325
guanylate cyclase, soluble, alpha chain - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S18325
R;Harteneck, C.; Wedel, B.; Koesling, D.; Malkewitz, J.; Boehme, E.; Schultz, G.
FEBS Lett. 292, 217-222, 1991
A;Title: Molecular cloning and expression of a new alpha-subunit of soluble guanylyl cyc
A;Reference number: S18325; MUID:92070494; PMID:1683630
A;Accession: S18325
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-732 <HAR>
A;Cross-references: EMBL:X63282; NID:g31670; PIDN:CAA44921.1; PID:g31671
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
F;473-701/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             45.5%;  Score 1635;  DB 2;  Length 732;
  Best Local Similarity   49.4%;  Pred. No. 7.9e-106;
  Matches  355;  Conservative  106;  Mismatches  200;  Indels   58;  Gaps   15;

Qy         13 GECPFSLLA-------PGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLPQRKTS 65

Db         27 GECPLSRLCWNGSRSPPG--PLEPSPAAAAAA---AAPAPTPAASAAAAAATAGARRVQR 81

Qy         66 RSRVYLHTLAESICKLIFPEFERLNVALQRTLA--KHKI---KESRKSL----------- 109

Db         82 RRRVNLDSLGESISRLTAPSPQTIQQTLKRTLQYYEHQVIGYRDAEKNFHNISNRCSYAD 141

Qy        110 -----EREDFEKTIAEQAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLN 164

Db        142 HSNKEEIEDVSGILQCTANILGLKFEEIQKRFGEEFFNICFHENERVLRAVGGTLQDFFN 201

Qy        165 SFSTLLKQSSHCQEA-GKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAA 223

Db        202 GFDALLE---HIRTSFGKQATLESPSFLCKELPEGTLMLHYFHPHHIVGFAMLGMIKAAG 258

Qy        224 HVLYETEVEVSLMPPCFHND--CSEFVNQP---YLLYSVHMKSTKPSLSPSKPQSSLVIP 278

Db        259 KKIYRLDVEVEQVA----NEKLCSD-VSNPGNCSCLTFLIKECENTNIMKNLPQGTSQVP 313

Qy        279 TSL------FCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKIN 332

Db        314 ADLRISINTFCRAFPFHLMFDPSMSVLQLGEGLRKQL-RCDTHKVLKFEDCFEIVSPKVN 372

Qy        333 QTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLED 392

Db        373 ATFERVLLRLSTPFVIRTKPEASGSENKDKVMEVKGQMIHVPESNSILFLGSPCVDKLDE 432

Qy        393 FTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVD 452

Db        433 LMGRGLHLSDIPIHDATRDVILVGEQAKAQDGLKKRMDKLKATLERTHQALEEEKKKTVD 492

Qy        453 LLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTR 512

Db        493 LLYSIFPGDVAQQLWQGQQVQARKFDDVTMLFSDIVGFTAICAQCTPMQVISMLNELYTR 552

Qy        513 FDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIK 572

Db        553 FDHQCGFLDIYKVETIGDAYCVAAGLHRKSLCHAKPIALMALKMMELSEEVLTPDGRPIQ 612

Qy        573 MRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGF 632

Db        613 MRIGIHSGSVLAGVVGVRMPRYCLFGNNVTLASKFESGSHPRRINVSPTTYQLLKREESF 672

Qy        633 VFTPRSREELPPNFPSEIPGICHFLDAYQQGTNSKPCFQKKDVEDGNAN----FLGKAS 687

Db        673 TFIPRSREELPDNFPKEIPGICYFLEVRTGPKPPKPSLSSSRIKKVSYNIGTMFLRETS 731


RESULT 5
JH0810
guanylate cyclase (EC 4.6.1.2) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 26-Aug-1999
C;Accession: JH0810
R;Yoshikawa, S.; Miyamoto, I.; Aruga, J.; Furuichi, T.; Okano, H.; Mikoshiba, K.
J. Neurochem. 60, 1570-1573, 1993
A;Title: Isolation of a Drosophila gene encoding a head-specific guanylyl cyclase.
A;Reference number: JH0810; MUID:93203896; PMID:8095978
A;Accession: JH0810
A;Molecule type: mRNA
A;Residues: 1-683 <YOS>
A;Cross-references: GB:S57126; NID:g298674; PID:g298675
A;Experimental source: head
C;Genetics:
A;Gene: dgc 1
A;Cross-references: FlyBase:FBgn0010227
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; phosphorus-oxygen lyase
F;417-641/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             27.2%;  Score 976.5;  DB 2;  Length 683;
  Best Local Similarity   37.3%;  Pred. No. 5.1e-60;
  Matches  244;  Conservative  116;  Mismatches  211;  Indels   83;  Gaps   20;

Qy         58 SLPQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAK--------HKIK---ESR 106

Db         27 ALEDEELSDDALTLTHLQMAIQLLTAPSNEDLNTAVTSLVAKYRQNWPNIHKLKLDPQTF 86

Qy        107 KSLEREDFEKTIAE------QAVAAGVPVEVIKESLGEEVFKICYEE-DENILGVVGGTL 159

Db         87 KSCANYDYLADIQELLLKMDEASASEILV-----LLGEELITCCCTGIIERAFRCLGTDL 141

Qy        160 KDFLNSFS---TLLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILP 216

Db        142 QEFLGSLDGVYDVLK----LQEED----VTDTGFVCAGEGE----LIFTSERPVIAWLLL 189

Qy        217 GIIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQPYLL----YSVHMKSTKPSLSPSKPQ 272

Db        190 GSLKALTRMLYKVDVNIKIEPV----ERCPAVSLPLLAGQGQLPDHADGSSTSVSKTIPE 245

Qy        273 ----------SSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEE 322

Db        246 TVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLY-------KPYMAD 298

Qy        323 -------YFEILTPK-INQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIYIV 374

Db        299 FGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIGLEIKGQMVHCP 358

Qy        375 ESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKA 434

Db        359 ESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQARAQDGLRRRMDKIKN 418

Qy        435 TLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAIC 494
```

```
Db        419 SIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYPDVTILFSDIVGFTSIC 478

Qy        495 SQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKES--DTHAVQIALM 552

Db        479 SRATPFMVISMLEGLYKDFDEFCDFFDVYKVETIGDAYCVASGLHRASIYDAHR---CLD 535

Qy        553 ALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612

Db        536 GLKMIDACSKHITHDGEQIKMRIGLHTGTVLAGVVGRKMPRYCLFGHSVTIANKFESGSE 595

Qy        613 PRKINVSPTTYRLLKDCPGFVF--TPRSREELPPNFPSEIPG---ICHFLDAYQ 661

Db        596 ALKINVSPTTKDWLTKHEGFEFELQPRDPSFLPKEFPN--PGGTETCYFLESFR 647


RESULT 6
OYRTB1
guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - rat
N;Alternate names: guanylate cyclase, soluble, 70K chain
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A31871
R;Nakane, M.; Saheki, S.; Kuno, T.; Ishii, K.; Murad, F.
Biochem. Biophys. Res. Commun. 157, 1139-1147, 1988
A;Title: Molecular cloning of a cDNA coding for 70 kilodalton subunit of soluble guanyla
A;Reference number: A31871; MUID:89087429; PMID:2905128
A;Accession: A31871
A;Molecule type: mRNA
A;Residues: 1-619 <NAK>
A;Cross-references: GB:M22562; NID:g204273; PIDN:AAA41204.1; PID:g204274
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             22.1%;  Score 795;  DB 1;  Length 619;
  Best Local Similarity   34.3%;  Pred. No. 1.9e-47;
  Matches  210;  Conservative  100;  Mismatches  225;  Indels   78;  Gaps   15;

Qy         89 LNVALQRTLAKHKIKESRKSLERE--------------DFEKT---IAEQAVAAGVPVE 130

Db          5 VNHALELLVIRNYGPEVWEDIKKEAQLDQEGQFLVRIIYDHSKTYDLVAAASKVLNLNAG 64

Qy        131 VIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDAS 189

Db         65 EILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDAL-----HDHLATIYPGMRAPS 119

Qy        190 ILCLDKE-DDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLM----PPCFHNDC 244

Db        120 FRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRSEECDHTQ- 178

Qy        245 SEFVNQPYLLYSVHMKSTKPS--------LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDM 296

Db        179 ----------FLIEEKESKEEDFYEDLDRFEENGTQDSRISPYT-FCKAFPFHIIFDRDL 227

Qy        297 TILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNS 356

Db        228 VVTQCGNAIYRVLPQLQ-PGKCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLD 286

Qy        357 VKK----------SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIH 406

Db        287 VEKLECEDELTGAEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346

Qy        407 NALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQL 466

Db        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406

Qy        467 WQGQVVQAKKFSNVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFDQQCGELD- 521

Db        407 RHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKN 466

Qy        522 --VYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHS 579

Db        467 PFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHT 525

Qy        580 GSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSR 639

Db        526 GEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL----------MSP 575

Qy        640 EELPPNFPSEIPG 652

Db        576 ENSDPQFHLEHRG 588


RESULT 7
OYBO70
guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - bovine
N;Alternate names: guanylate cyclase, soluble, 70K chain
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
C;Accession: S01653; A38759
R;Koesling, D.; Herz, J.; Gausepohl, H.; Niroomand, F.; Hinsch, K.D.; Muelsch, A.; Bo
FEBS Lett. 239, 29-34, 1988
A;Title: The primary structure of the 70 kDa subunit of bovine soluble guanylate cycl
A;Reference number: S01653; MUID:89031214; PMID:2903071
A;Accession: S01653
A;Molecule type: mRNA
A;Residues: 1-619 <KOE>
A;Cross-references: EMBL:Y00770; NID:g407; PIDN:CAA68739.1; PID:g408
A;Accession: A38759
A;Molecule type: protein
A;Residues: 1-6;28-32;41-45;96-109;337-343;396-406;562-569 <KOE2>
C;Genetics:
A;Introns: 281/3; 392/2
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             22.1%;  Score 793.5;  DB 1;  Length 619;
  Best Local Similarity   35.3%;  Pred. No. 2.5e-47;
  Matches  208;  Conservative   97;  Mismatches  217;  Indels   67;  Gaps   16;

Qy        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153

Db         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88

Qy        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212

Db         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDADKGKGLILHYYSEREGLQ 143

Qy        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265

Db        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192

Qy        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320

Db        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250

Qy        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKSS----------RVMDLKGQM 370

Db        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKSECEDELTGTEISCLRLKGQM 310

Qy        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430

Db        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370

Qy        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490

Db        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430

Qy        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543

Db        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490

Qy        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603
```

```
Db        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549

Qy        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652
              :: |:     ||||  ||| |        ||   |   | | |  |
Db        550 TSRTETTGEKGKINVSEYTYRCL-------MTP---ENSDPQFHLEHRG 588


RESULT 8
S23097
guanylate cyclase (EC 4.6.1.2), soluble, 70K chain - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 20-Apr-2000
C;Accession: S23097
R;Giuili, G.; Scholl, U.; Bulle, F.; Guellaen, G.
FEBS Lett. 304, 83-88, 1992
A;Title: Molecular cloning of the cDNAs coding for the two subunits of soluble guanylyl
A;Reference number: S23097; MUID:92316204; PMID:1352257
A;Accession: S23097
A;Molecule type: mRNA
A;Residues: 1-619 <GIU>
A;Cross-references: EMBL:X66533; NID:g31685; PIDN:CAA47144.1; PID:g31686
C;Genetics:
A;Gene: GDB:GUCY1B3; GUC1B3; GC-SB3
A;Cross-references: GDB:141992; OMIM:139397
A;Map position: 4q31.3-4q33
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: heterodimer; phosphorus-oxygen lyase
F;373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             22.0%;  Score 789.5;  DB 1;  Length 619;
  Best Local Similarity   35.1%;  Pred. No. 4.7e-47;
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153
              :: |  : | |  :: :     |||   |      |:: :  |:  |  | |   : ||
Db         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88

Qy        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212
              |:|  :::|| :   |      |   |      :   |  | | |    | ::|:  :
Db         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143

Qy        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265
              |: ||||  |  :: ||::: ::      | |             : :  | :|
Db        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192

Qy        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
                  :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250

Qy        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370
              | :: | |: :| ||:: :|  ||:| :     |:|                : |||||
Db        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310

Qy        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430
              ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370

Qy        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490
              |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430

Qy        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543
              | ||: :      :::: :|| ||||||          ||||||:|| |    || :
Db        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490

Qy        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603
              ||  |  :|| |||:: :|    || ::: ||:|:|  |  ||:| :||||||||| | |
Db        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549

Qy        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652
              :: |:     ||||  ||| |          | |   | |  |  |
Db        550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588


RESULT 9
OYRTB2
guanylate cyclase (EC 4.6.1.2), soluble, beta-2 chain - rat
N;Alternate names: guanylate cyclase, soluble, 76K chain
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A36228
R;Yuen, P.S.T.; Potter, L.R.; Garbers, D.L.
Biochemistry 29, 10872-10878, 1990
A;Title: A new form of guanylyl cyclase is preferentially expressed in rat kidney.
A;Reference number: A36228; MUID:91105012; PMID:1980215
A;Accession: A36228
A;Molecule type: mRNA
A;Residues: 1-682 <YUE>
A;Cross-references: GB:M57507; GB:J05308; NID:g204279; PIDN:AAA41207.1; PID:g204280
A;Experimental source: kidney
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;360-584/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             21.6%;  Score 774.5;  DB 1;  Length 682;
  Best Local Similarity   32.6%;  Pred. No. 5.9e-46;
  Matches  190;  Conservative  107;  Mismatches  204;  Indels   81;  Gaps   10;

Qy        129 VEVIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLED 187
              :| | :  ||  || |     : :|  :|| | :|: :   |     |   |     :
Db          1 MEAILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDAL-----HSYLALSYQEMNA 55

Qy        188 ASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCFHNDCSEF 247
              |    :  |  : ::|:  :     |:||||:| |   ::|:| :|::      | :|
Db         56 PSFRVEEGADGAMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMSIL------DMNEE 109

Qy        248 V----NQPYLLYSVHMKSTKP--SLSPSKPQSS------------------------- 274
              |     : :::: |  |: :        |:|| |
Db        110 VERTGKKEHVVFLVVQKAHRQIRGAKASRPQGSEDSQADQEALQGTLLRMKERYLNIPVC 169

Qy        275 -------------------------------LVIPTSLFCKTFPFHFMFDKDMTILQFG 302
                                            | :   :||  |||| :||: : : | |
Db        170 PGEKSHSTAVRASVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFHIVFDEALRVKQAG 229

Qy        303 NGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRR-WDNSVKKSS 361
                |:: :     | |   :||| |: |::    | |   :| |||:: |:      :||
Db        230 VNIQKYVPGILTQ-KFALDEYFSIIHPQVTFNISSICKFINSQFVLKTRKEMMPKARKSQ 288

Qy        362 RVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARA 421
              :: |:||||::     ::|: || |  |::     ::|||| |:  ||::|: :|  |
Db        289 PMLKLRGQMIWMESLRCMIFMCSPNVRSLQELEESKMHLSDIAPHDTTRDLILLNQQRLA 348

Qy        422 QDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVT 481
              :  |  :| | |  |       |  ||||  || :: |  || || :|: | | :|   |
Db        349 EMELSCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVANQLKEGRKVAAGEFETCT 408

Qy        482 MLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKE 541
              :||||:| || ||: | |:|:: |||::|:::||:     ||||||||||| | ||:
Db        409 ILFSDVVTFTNICAACEPIQIVNMLNSMYSKFDRLTSVHDVYKVETIGDAYMVVGGVPVP 468

Qy        542 SDTHAVQIALMALKMMELSDEVMSP-HGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ::|| ::|  || |    : |||:|  ||||::|:|:|:| | ||||| |||||||||:
Db        469 VESHAQRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVVGDKMPRYCLFGDT 528

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              |  |:: ||  :| |:::|||  :| ||:  ||    |   |:
Db        529 VNTASRMESHGLPSKVHLSPTAHRALKN-KGFEIVRRGEIEV 569


RESULT 10
T24458
hypothetical protein T04D3.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T24458
R;Kershaw, J.
```

```
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19893
A;Accession: T24458
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-751 <WIL>
A;Cross-references: EMBL:Z81114; PIDN:CAB03288.1; GSPDB:GN00019; CESP:T04D3.4
A;Experimental source: clone T04D3
C;Genetics:
A;Gene: CESP:T04D3.4
A;Map position: 1
A;Introns: 1/3; 26/2; 95/3; 149/1; 213/3; 247/3; 291/3; 332/3; 369/2; 418/1; 462/3; 498/
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match             20.9%;  Score 749.5;  DB 2;  Length 751;
  Best Local Similarity   31.5%;  Pred. No. 3.7e-44;
  Matches  192;  Conservative  125;  Mismatches  239;  Indels   53;  Gaps   16;

Qy        116 KTIAEQAVAAGVPVEVIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSS 174
Db         50 RLVNSMANVIGIPIEEIWEAYGGFLIQFTMETGWDELLRAMAPDLEGFLDSLDSLHYFID 109

Qy        175 HCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVS 234
Db        110 H---VVYKTKLRGPSFRCDVQADGTLLLHYYSKRSGLYPIVKGVVREVARRIYDTEVVMK 166

Qy        235 LMP-PCFHNDCSEFVNQPYLLYSV----HMKSTKPSLSPSKPQSSLVIPTSL-------F 282
Db        167 VQERKQEHLDA--FVTE-HVVFVITQIENANSTQPKSISSKADSQIDLSTGIYEISSSDF 223

Qy        283 CKTFPPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTML 342
Db        224 SLAFPYHICFDPDLFVEHFGNFIKKTFPNA-MRQETRVTDLLELVHPEVPFSYESIKYYK 282

Qy        343 NMQFVVRVRRWDNSVKKS---SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLY 399
Db        283 NSLFVFRLKGLGDIVHNANDEAKTVLLKGSMVFIDEGKYILYMCSVNVTTVRELIERNLH 342

Qy        400 LSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFP 459
Db        343 LSDMQRHDGTRDVIMLNQSRMSQVELNRTLEETTKKLKKMAQELEIEKQKTDELLCELMP 402

Qy        460 CEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGE 519
Db        403 ASVADSLRSGKAMDAKEFADCTLLFTDIVTFTNICAMCTPYDVVTLLNDLYLRFDRLVGL 462

Qy        520 LDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSP--HGEPIKMRIGL 577
Db        463 HDAYKVETIGDAYMIVGGVPERCENHAERVLNISIGMLMESKLVLSPITH-KPIKIRLGV 521

Qy        578 HSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPT-TYRLLKDCPGFVFTP 636
Db        522 HCGPVVAGVVGIKMPRYCLFGDTVNVANKMESNGIQCKIHVSETGKLNGLKANPSYVFID 581

Qy        637 RSREE-----------LPPNFPSEIPGIC-------HFLDAYQQGTNSKPCFQKK----- 673
Db        582 RGNTEIRGKGMMYTYFLERNDRKSVWELCSRPRSGEQTIDGYME-LHDQSIYQEEGGQQE 640

Qy        674 --DVEDGNA 680
Db        641 NLTVENGNS 649


RESULT 11
T19968
hypothetical protein C46E1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T19968; T23092
R;Lennard, N.
submitted to the EMBL Data Library, November 1997
A;Reference number: Z19205
A;Accession: T19968
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-685 <WIL>
A;Cross-references: EMBL:AL008867; PIDN:CAA15512.1; GSPDB:GN00028; CESP:C46E1.2
A;Experimental source: clone C46E1
R;Lennard, N.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z19673
A;Accession: T23092
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-685 <WI2>
A;Cross-references: EMBL:Z99942; PIDN:CAB17073.1; GSPDB:GN00028; CESP:C46E1.2
A;Experimental source: clone H13N06
C;Genetics:
A;Gene: CESP:C46E1.2
A;Map position: X
A;Introns: 11/3; 36/2; 70/1; 133/2; 179/1; 202/3; 258/2; 298/2; 301/3; 325/2; 405/2;
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match             16.3%;  Score 587;  DB 2;  Length 685;
  Best Local Similarity   28.1%;  Pred. No. 6.8e-33;
  Matches  185;  Conservative  114;  Mismatches  254;  Indels  106;  Gaps   20;

Qy        107 KSLEREDFEKTIAEQAV-------------AAGVPVEVIKESLGE--EVFKICYEED-- 148
Db         34 KVLERAGFEAG-KENIINHYYSDADTFSLVDAVSVILKVTREQVWEMYGCFLIQYTMETG 92

Qy        149 -ENILGVVGGTLKDFLNSFSTLLKQSSH-CQEAGKRGRLEDASILCLDKEDDFLHVYYFF 206
Db         93 WDDLIRSMSPNLKGFLDNLDSLHYFIDHVVYKANLRG----PSFRCEDNPDGTITLHYYT 148

Qy        207 PKRTTSLILPGIIKAAAHVLYETEVEVSL---MPPCFHNDCSEFVNQPYLLYSVHM---- 259
Db        149 GRPGLYPIVKGVLREAAKRVFKLDVSMTITGRTQRSVQMATGERIEE-HVIFLVKTLNTD 207

Qy        260 KSTKPSLSPSKPQSSLVIPTSL----FCKTFPPFHFMFDKDMTILQFGNGIRRLMNR--RD 313
Db        208 QSNEEALGTAVVQHSNNYKIRLTHMDFISTFPYHMVVDQDCKIVQVG---RELYNHIPKD 264

Qy        314 F--QGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWD-----NSVKKSSRVMD- 365
Db        265 LLSVGTP-LMRIFEVTRPQIPLDFDSICNFINAVFVLQVKTTPMEFQRNANKRAAQAIEA 323

Qy        366 --------------------LKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIP 404
Db        324 SENLYEDNNGALALSQSQHLKLKGQMMLMSSGGHIMYLCSPYVTSIPELLQYGLRLTAMP 383

Qy        405 IHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQ 464
Db        384 IHDPTRDLILLNQQRLSDVEMNLQLEANNEQLENMAKDLEVEKGKTDALLREMLPPSVAQ 443

Qy        465 QLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYK 524
Db        444 QLKQGLSVEAREYEEATVMFTDVPTFQQIVPLCTPKDIVHLLNELFTKFDRLIGIQKAYK 503

Qy        525 VETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSP-HGEPIKMRIGLHSGSVF 583
Db        504 VETVGDSYMSVGGIPDLVDDHCEVICHLALGMVMEARTVCDPITNTPLHIRAGIHSGPVV 563

Qy        584 AGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREE-- 641
Db        564 AGVVGAKMPRYCLFGDTVNTSSRMESHSPIGRIHCSENAKKCAESTGRFEFEPRGRVQIK 623

Qy        642 ----------LPPNFPSEIPGI--------CHFLDAYQQ-------------GTNSKPC 669
Db        624 GKGEMNTYFLLRSFKRSIWEIIDRRRDENCNSIDGYNELREGYVDDVLNKVTQKNSKTC 682


RESULT 12
T23721
hypothetical protein M04G12.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
```

```
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T23721
R;Sims, M.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19789
A;Accession: T23721
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-686 <WIL>
A;Cross-references: EMBL:Z81103; PIDN:CAB03210.1; GSPDB:GN00023; CESP:M04G12.3
A;Experimental source: clone M04G12
C;Genetics:
A;Gene: CESP:M04G12.3
A;Map position: 5
A;Introns: 1/3; 26/2; 60/1; 192/3; 255/2; 321/2; 401/2; 645/1
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match             16.2%;  Score 582.5;  DB 2;  Length 686;
  Best Local Similarity   27.2%;  Pred. No. 1.4e-32;
  Matches  169;  Conservative  126;  Mismatches  265;  Indels   61;  Gaps   16;

Qy          70 YLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAEQAV-AAGVP 128
               ::|   ||| :|:  ::   :| ||  | :   :  ::::   :  |    | :  :
Db           4 FIH---ESIRQLVIRKYGE-DVWLQ-VLERSGFENGKENIVNHYYSDTDTYVLVDSVSIV 58

Qy         129 VEVIKESLGEEV--FKICYEED---ENILGVVGGTLKDFLNSFSTLLKQSSH-CQEAGKR 182
               ::| |: : |    | | |  :   : ::  :   || ||::  :|     |   :|  |
Db          59 LKVTKDQIWEMYGGFLITYSMEIGWDELVRSMSPNLKGFLDNLDSLHYFIDHVVYKANLR 118

Qy         183 GRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLM---PPC 239
               |     | | :  |  | ::||  :     |: |::|  | :::  :: : :
Db         119 G----PSFRCEENPDGTLMLHYFTGRPGLYHIVKGVVKEVAKLVFNLDISLVVQGRTQRS 174

Qy         240 FHNDCSEFVNQP--YLLYSVHM------KSTKPSLSPSKPQSSLVIPTSL------FCKT 285
               | :  | | :   :|: :|        ||  :|:  :|    :|  :|      | :
Db         175 VHMNNGERVEEHVIFLIKNVEEPRRDSDTSTTSALTSVEPDFGEIIDDNLKVSLQDFSRA 234

Qy         286 FPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQ 345
               |:||: |:   ::| |: :   :     |        |||  |:|   |  |   :|
Db         235 LPYHFVLDESCRLVQCGDELYNHIPNELLQPGTPILRIFEINRPQIPLDFENICNFINAV 294

Qy         346 FVVRV------RRWDNSVKKSSRVMD-------------------LKGQMIYIVESSAI 379
               ||:|      ::  |:: |  |  :                    |||||: :     |
Db         295 FVLQVKTSPLRKKHMNAMTKEEREQEVEAMEEEVESNELTQGCHLKLKGQMMMLSTKKHI 354

Qy         380 LFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQA 439
               ::| || |  : :    |: |: :|:|:| ||::|: :|  :   :  :|      ||
Db         355 IYLCSPYVTSINELMQFGMRLTAMPLHDATRDLILLNQQRLSDVEVNLQLEANNEQLETM 414

Qy         440 HQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSP 499
                  || |::||  :|  : | ::|:||  |: ::  ::   |::| |:  |  |   | |
Db         415 THELEVERQKTDSILKDMLPRKIAKQLLSGEHLEPCEY-EATVMFCDLPAFQQIIPVCQP 473

Qy         500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMEL 559
                 :: :|| :: : |:      ||||||: |:|    |:   :  ||  :  :|| ||
Db         474 KNIVKLLNEVFFKLDRIVVLRGVYKVETVSDSYMTVSGIPDYTSEHAENMCHVALGMMWE 533

Qy         560 SDEVMSP-HGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINV 618
               :  || | :  |  :|||||||:: ||||| |||||||||  ||||:: ||  |  ||
Db         534 ARSVMDPVNKTPFLLRIGLHSGTIIAGVVGTKMPRYCLFGETVTLASQMESLGVAGKIQC 593

Qy         619 SPTTYRLLKDCPGFVFTPRSR 639
               |  ||    :   | |:|| |
Db         594 SSWTYSKAMETGRFEFSPRGR 614


RESULT 13
T16822
hypothetical protein T07D1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16822
R;Leimbac, D.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid T07D1.
A;Reference number: Z18584
A;Accession: T16822
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1099 <LEI>
A;Cross-references: EMBL:U41531; NID:g1109802; PID:g1109803; PIDN:AAA83158.1; CESP:T0
C;Genetics:
A;Gene: CESP:T07D1.1
A;Introns: 26/3; 94/3; 144/2; 200/3; 239/2; 278/3; 344/1; 382/2; 425/3; 474/3; 523/3;

  Query Match             16.2%;  Score 581;  DB 2;  Length 1099;
  Best Local Similarity   28.4%;  Pred. No. 3.4e-32;
  Matches  177;  Conservative   97;  Mismatches  204;  Indels  146;  Gaps   19;

Qy         126 GVPVEVIKESLGEEVFKICYEEDEN----------ILGVVGGTLKDFLN---------- 164
               | ||: :  ::|   :|   : : |           :| |:| |   |||
Db          59 GDPVDELMTNIGTSFYKFLTKFEFNKVYYVLFYIPYVLRVLGRTFPQFLNGLDNLHEYLR 118

Qy         165 -SFSTLLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAA 223
                :|  |   | :|:   : |     ::    |  |||           : | |: :
Db         119 FTFPKLKPPSFYCEHESRTG----LTLHYRSKRRGFLH-----------YVQGQIRNIS 162

Qy         224 HVLYETEVEVSLMPPCFHNDCSEFVNQPYLLYSVHMKS---TKPSLSPSKPQSSLV---- 276
               |::||| : |:      |    :|  :::  :|  :    :   :      |::
Db         163 QELFQTEVVIELL------DIEHDLNLEHVIMRLHFNNLDFNRQGTAYRNLNDSILEKVK 216

Qy         277 IPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLM-------------------------- 309
               | : :|   |||  :|:: | |   | |: |:|
Db         217 ITSDIFFDIFPFIIVFNRGMRIRNIGIGLLRVMAGIVGKKINQTFLLMRPFIRFRWEEVR 276

Qy         310 --NR-----------RDFQGKP--------------NFEEYFEIL-TPKINQTFSGIM-- 339
                 ||           |:|   |              :   ||:: :  | :   ||:
Db         277 GTNRNYSFDIACDCFRNFFASPKKRVLKSVKSQIMLHSNNIFELISSDPIQEDEDGILVY 336

Qy         340 --TMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRG 397
                 | ::    | |  |   :|    : ||||| |: |  :| |:| | :  |      |
Db         337 KTTDVDQMTEERHRMGDGEREK---FLSLKGQMFYMEEWESICFVGIPVMSHLPQMYKSG 393

Qy         398 LYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQ----------ALEEEK 447
               |:::|  :|:: ||:||   |  |:          ||  | |  |           |::|:
Db         394 LFINDFALHDSSRDLVLASTQQSAE---------LKLLLHQEAQKSRNMRENMNRLKKER 444

Qy         448 KKTVDLLCSIFPCEVAQQLWQGQ--VVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITM 505
               ::|  ||  : |  || ||  |:  |   ::| :||:||:||| || :||  :||:||
Db         445 RRTDKLLYQMLPKSVANQLRHGESAVACCERFDSVTILFTDIVEFTKMCSSLTPLEVIEF 504

Qy         506 LNALYTRFDQQCGELDVYKVETIGDAYCVAGG--LHKESDTH-----AVQIALMALKMME 558
               |  :|| ||:       ||||||||||| |  |     | |       | |  : | ||:
Db         505 LKVIYTNFDKIIDTHGVYKVETIGDAYMVVSGAPTKTEHDAEFILDCASQFLVEAGKMVN 564

Qy         559 LSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINV 618
               :::::       | :| |:|||| |||||: ||||||||  | :||| |  | | || |
Db         565 MNNKI-----HKIDIRAGVHSGSVVAGVVGLSMPRYCLFGETVYVANKMEQNSSPMKILV 619

Qy         619 SPTTYRLLKDCPGFVFTPRSREEL 642
               | ||:  :::    ::    |||:
Db         620 SETTHNKIEESDPGLYQFERREEI 643


RESULT 14
T18984
hypothetical protein C06B3.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T18984
R;Percy, C.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19056
A;Accession: T18984
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
```

```
A;Residues: 1-699 <WIL>
A;Cross-references: EMBL:Z77652; PIDN:CAB01118.1; GSPDB:GN00023; CESP:C06B3.8
A;Experimental source: clone C06B3
C;Genetics:
A;Gene: CESP:C06B3.8
A;Map position: 5
A;Introns: 1/3; 60/1; 192/3; 210/2; 270/2; 336/2; 415/2; 659/1
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match            15.3%;  Score 551.5;  DB 2;  Length 699;
  Best Local Similarity   26.6%;  Pred. No. 2.1e-30;
  Matches  153;  Conservative  112;  Mismatches  240;  Indels   71;  Gaps   12;

Qy        129 VEVIKESLGEEV--FKICYEED---ENILGVVGGTLKDFLNSFSTLLKQSSH-CQEAGKR 182
              ::| |: : |    | | |  :   : ::  :   || ||::  :|     |   :|  |
Db         59 LKVTKDQVWEMYGGFLITYSMEIGWDELVRSMSPNLKGFLDNLDSLHYFIDHVVYKANLR 118

Qy        183 GRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLM---PPC 239
              |     | | :  |  | ::||  :     |: |::|  |  ::: :: : :
Db        119 G----PSFRCEETPDGTLLLHYFTGRPGLYHIVKGVVKEVAKRVFDLDITLVVQGRTQRS 174

Qy        240 FHNDCSEFVNQ-------------PYLLYSVHMKSTKPSLSPSKPQSSLVIPTSL---- 281
              | :  | | :               |:|: :: :  ::|       : ||:  |:
Db        175 VHMNNGERVEEHVVFLINKKGGQFHPFLI-AIVVTGSEPRRDSEGSEVSLLTSTNANFPT 233

Qy        282 ------------FCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTP 329
                         | |  |:||: |:   ::| |: :   :      |         ||| |
Db        234 IVDDTLGISLDDFSKALPYHFVIDESCKLVQCGSELHNHIPNELLQPGTPILRIFEINRP 293

Qy        330 KINQTFSGIMTMLNMQFVVRVR-----------RWDNSVKKSSRVMD------------ 365
              :|   | |   :|  ||:|:                 :|:    :|
Db        294 QIPLDFENICNFINAVFVLQVKTSPLKKKHMDAMSQEELKQEMETLDEDATNELTQGHHL 353

Qy        366 -LKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDG 424
               |||||: :     |::| || |  : :    |: |: :|:|:| ||::|: :|  :
Db        354 KLKGQMMLLASKKHIIYLCSPYVTSINELMQYGMRLTAMPLHDATRDLILLNQQRLSDVE 413

Qy        425 LKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLF 484
              :  :|      ||  : || |::||  :|  :.|  :||||  |: ::| :    |::|
Db        414 VNLQLEANNEQLETMTRELELERQKTDSILKDMLPRRIAQQLLSGEHIEACE-HEATVMF 472

Qy        485 SDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDT 544
              |: |     ||||  :: ||| :: : |:      ||||||: |:|    |:   :
Db        473 CDLPAFQQAIPQCSPKDIVNMLNEIFRKLDRIVVIRGVYKVETVSDSYMAVSGIPDYTPE 532

Qy        545 HAVQIALMALKMMELSDEVMSPHGE-PIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603
              ||  :  :|| ||  :  |: |  : |  :|||:|||:: |||||   |:|||||  |||
Db        533 HAENMCHVALGMMWEARSVIDPVSKTPFLLRIGIHSGTITAGVVGTVHPKYCLFGETVTL 592

Qy        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSR 639
              |:: ||  :  ||  |  |:   :   | |:|| |
Db        593 ASQMESLGMAGKIQCSKWAYQKAMETGRFEFSPRGR 628


RESULT 15
F88642
protein C54E4.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: F88642
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: F88642
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-583 <STO>
A;Cross-references: GB:chr_IV; PIDN:AAB92031.1; PID:g2702405; GSPDB:GN00022; CESP:C54E4.
A;Note: Similar to guanylate cyclase
C;Genetics:
A;Gene: C54E4.3
A;Map position: 4

  Query Match            13.7%;  Score 493.5;  DB 2;  Length 583;
  Best Local Similarity   25.7%;  Pred. No. 1.8e-26;
  Matches  139;  Conservative  129;  Mismatches  218;  Indels   55;  Gaps   13;

Qy        122 AVAAGVPVEVIKESLGE-EVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              |   |: |: : |  ||  :   |   : :|  :   |::||:: :::       :
Db         50 ATVLGLSVDDMWEMYGEFLITHACETGWQKMLFCMANNLQEFLDNLNSM---HYFIDQIA 106

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              :  ::  :  |    :  | ::||  ::    |: |:::  |  |:| :|:| ::
Db        107 FKSEMKGPTFQCEPFGESGLKLHYFSFRQGLFPIVKGLVRKTARTLFEMDVKVCMLERNQ 166

Qy        241 HNDCSEFVNQPYLLYSVH-------------MKSTK-----PSLSPSKPQSSLVIPTSL 281
              |  |   ::::||                 ::||      || : |  |::::
Db        167 ERRKSGMVE--HVIFSVEPDDNHRKGKRLFHKFRNTKTTENAPSFTLS---STILVGLRD 221

Qy        282 FCKTFPFHFMFDKDMTILQFGNGIRR---LMNRRDFQGKPNFEEYFEILTPK-INQTFSG 337
              |   ||:|  |:| | |   |  : |   | |::  :      :  ::: |  |  |:
Db        222 FKNIFPYHVCFNKQMIIEHIGIYLLREYGLENKKTLK----VSDLMQLVQPSDIQLTYKN 277

Qy        338 IMTMLNMQFVVRVRRWD--NSVKKSS-----RVMDLKGQMIYIVESSAILFLGSPCVDRL 390
              ::: ||  |: :::     | |:: |     : : |||:|: | : ::|:|: || |  :
Db        278 VLSYLNTLFIFQLKHHSKRNEVQEGSSEAFQQPLVLKGEMMPINDGNSIIFICSPHVTTV 337

Qy        391 EDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKT 450
               |    ||:||:|:|:| ||:|:: :    |  | |:| :   :::  : || :| :|
Db        338 RDILNLKLYISDMPMHDATRDLVMLNQSRICQMELNKKLEETMKKMKKMTEELEVKKSQT 397

Qy        451 VDLLCSIFPCEVAQQLWQGQVVQA----------KKFSNVTMLFSDIVGFTAICSQCSP 499
               ||    |  :|: |   : | |           ::||: :::|:||  |  |   |||
Db        398 DRLLFEFVPPVIAEALRAAKTVPALMQKRISIDFSEEFSDCSVIFTDIPDFFTISVNCSP 457

Qy        500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMEL 559
              ::||::  |: |||:   :   ||| :: |:| : ||:   :  |      :|| ::
Db        458 TEIITVVTDLFHRFDRIIEKHKGYKVLSLMDSYLIVGGVPNANQYHCEDSLNLALGLLFE 517

Qy        560 SDEVMSPHGE-PIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINV 618
              : :|: |  |  :::|||:| | | ||:|  : ||:|: || | :     | | | |: |
Db        518 AKQVVVPKLERSVRLRIGVHCGPVVAGIVSQQKPRFCVLGNTVNVTKSICSHSSPGKVLV 577

Qy        619 S 619
              |
Db        578 S 578


Search completed: June 27, 2003, 13:05:23
Job time : 21.5034 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:00:58 ; Search time 34.7899 Seconds
(without alignments)
4086.601 Million cell updates/sec

Title: US-09-762-767A-2
Perfect score: 3593
Sequence: 1 MFCTKLKDLKITGECPFSLL.....QKKDVEDGNANFLGKASGID 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3587	99.8	690	4	Q8TAH3	Q8tah3 homo sapien
2	3266.5	90.9	691	11	Q9DBQ3	Q9dbq3 mus musculu
3	3265.5	90.9	691	11	Q9ERL9	Q9erl9 mus musculu
4	2954	82.2	564	4	Q9NNW8	Q9nnw8 homo sapien
5	2014.5	56.1	678	13	Q9PWI2	Q9pwi2 oryzias lat
6	2009.5	55.9	678	13	P79997	P79997 oryzias lat
7	2005	55.8	675	13	Q90VV5	Q90vv5 fugu rubrip
8	1097	30.5	699	5	O77105	O77105 manduca sex
9	991.5	27.6	676	5	Q24085	Q24085 drosophila
10	985.5	27.4	676	5	Q95SQ4	Q95sq4 drosophila
11	823	22.9	220	11	Q8R5L4	Q8r514 rattus norv
12	788	21.9	617	13	Q90VY5	Q90vy5 fugu rubrip
13	785.5	21.9	600	5	O77106	O77106 manduca sex
14	784.5	21.8	742	11	Q91XJ7	Q91xj7 rattus norv
15	781.5	21.8	620	11	O54865	O54865 mus musculu
16	781	21.7	614	13	P79998	P79998 oryzias lat
17	771.5	21.5	636	11	Q920Q1	Q920q1 rattus norv
18	750.5	20.9	604	5	Q95NK5	Q95nk5 hemicentrot
19	749.5	20.9	751	5	O02298	O02298 caenorhabdi
20	743	20.7	758	5	Q24086	Q24086 drosophila
21	743	20.7	787	5	Q9VA09	Q9va09 drosophila
22	720.5	20.1	649	5	Q17010	Q17010 anopheles g
23	661	18.4	940	5	O76340	O76340 manduca sex
24	642.5	17.9	690	5	Q9VFC5	Q9vfc5 drosophila
25	587	16.3	685	5	Q9XTE0	Q9xte0 caenorhabdi
26	582.5	16.2	686	5	P92006	P92006 caenorhabdi
27	555	15.4	752	5	Q22301	Q22301 caenorhabdi
28	551.5	15.3	699	5	Q17707	Q17707 caenorhabdi
29	547	15.2	684	5	Q9BI80	Q9bi80 caenorhabdi
30	493.5	13.7	583	5	O44468	O44468 caenorhabdi
31	479	13.3	1047	13	Q9PWH0	Q9pwh0 xenopus lae
32	478	13.3	1276	5	Q8T8L6	Q8t816 caenorhabdi
33	477	13.3	1055	13	Q9YGW3	Q9ygw3 oryzias lat
34	473	13.2	1055	13	Q98UI1	Q98ui1 oryzias lat
35	463.5	12.9	1057	5	Q18331	Q18331 caenorhabdi
36	462	12.9	1076	5	Q9W2P1	Q9w2p1 drosophila
37	461	12.8	1112	5	Q9GT36	Q9gt36 heterodera
38	460	12.8	344	11	Q8R1P8	Q8r1p8 mus musculu
39	460	12.8	618	4	Q8TA93	Q8ta93 homo sapien
40	456.5	12.7	1172	5	Q9VU79	Q9vu79 drosophila
41	455.5	12.7	1056	13	Q9YI17	Q9yi17 squalus aca
42	454.5	12.6	1057	11	Q91X04	Q91x04 mus musculu
43	453.5	12.6	210	13	Q9NH71	Q9nh71 eptatretus
44	453	12.6	1076	5	Q9VF19	Q9vf19 drosophila
45	452	12.6	1100	11	O54884	O54884 rattus norv

ALIGNMENTS

RESULT 1
Q8TAH3
ID Q8TAH3 PRELIMINARY; PRT; 690 AA.
AC Q8TAH3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to guanylate cyclase 1, soluble, alpha 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC028384; AAH28384.1; -.
SQ SEQUENCE 690 AA; 77470 MW; D9B98FC0E4DBB927 CRC64;

Query Match 99.8%; Score 3587; DB 4; Length 690;
Best Local Similarity 99.7%; Pred. No. 1.6e-272;
Matches 688; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATMPICQDIPEKNIQESLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
```

```
Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMAVKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660

Qy        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
              |||||||||||||||||||||||||||||
Db        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
```

```
RESULT 2
Q9DBQ3
ID   Q9DBQ3      PRELIMINARY;      PRT;   691 AA.
AC   Q9DBQ3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   1200016O07Rik protein.
GN   GUCY1A3 OR 1200016O07RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=LUNG;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK004815; BAB23586.1; -.
DR   HSSP; P19687; 1AWN.
DR   MGD; MGI:1926562; Gucy1a3.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   691 AA;  77587 MW;  FF8DD0C3CB0BBD7F CRC64;

  Query Match             90.9%;  Score 3266.5;  DB 11;  Length 691;
  Best Local Similarity   89.7%;  Pred. No. 2.1e-247;
  Matches  620;  Conservative   36;  Mismatches   34;  Indels    1;  Gaps    1;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||| | |||||||||||||||||| | :|| || || |:||:|||| |||| : |||
Db          1 MFCRKFKDLKITGECPFSLLAPGQVPKEPTEEVAGGSEGCQATLPICQYFPEKNAEGSLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||:|||||||||||||||| ||||:||||||||||:|:||| |:|| || |||
Db         61 QRKTSRNRVYLHTLAESICKLIFPECERLNLALQRTLAKHKIEENRKSSEKEDLEKIIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              :|:||| ||| :|:|||||:||||||||:|||||||||||||||||||||||||||||||
Db        121 EAIAAGAPVEALKDSLGEELFKICYEEDEHILGVVGGTLKDFLNSFSTLLKQSSHCQEAE 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              :|||||||||||||: |||:||||||||||:|:|||||||| :|||: ||||||||||
Db        181 RRGRLEDASILCLDKDQDFLNVYYFFPKRTTALLLPGIIKAAARILYESHVEVSLMPPCF 240

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              :|||:|||||||||||:||||||||| |||||||||| ||||||||||| |:|: |||
Db        241 RSDCTEFVNQPYLLYSVHVKSTKPSLSPGKPQSSLVIPASLFCKTFPFHFMLDRDLAILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              |||||||:|:||||||||||:|||||||||||||||||||||||||:|||||||||||||
Db        301 LGNGIRRLVNKRDFQGKPNFEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              |||||||||||||||||||||||||:|||||||||| |||||||||:||||||| |
Db        421 AQDGLKKRLGKLKATLEHAHQALEEEKKRTVDLLCSIFPSEVAQQLWQGQIVQAKKFSEV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHR 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLDAY 660

Qy        661 -QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
               || |||| || ||||||||||||||||||
Db        661 HHQGPNSKPWFQDKDVEDGNANFLGKASGID 691
```

```
RESULT 3
Q9ERL9
ID   Q9ERL9      PRELIMINARY;      PRT;   691 AA.
AC   Q9ERL9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase alpha 1 subunit.
GN   GUCY1A3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE=20461464; PubMed=10984516;
RA   Sharina I.G., Krumenacker J.S., Martin E., Murad F.;
RT   "Genomic organization of alpha 1 and beta 1 subunits of the mammalian
RT   soluble guanylyl cyclase genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10878-10883(2000).
DR   EMBL; AF297082; AAG17446.1; -.
DR   HSSP; P19687; 1AWN.
DR   MGD; MGI:1926562; Gucy1a3.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   691 AA;  77573 MW;  FF9F40C3CB0BBD7F CRC64;

  Query Match             90.9%;  Score 3265.5;  DB 11;  Length 691;
  Best Local Similarity   89.6%;  Pred. No. 2.5e-247;
  Matches  619;  Conservative   37;  Mismatches   34;  Indels    1;  Gaps    1;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||| | |||||||||||||||||||| | :|| || || |:||:|||| |||| : |||
Db          1 MFCRKFKDLKITGECPFSLLAPGQVPKEPTEEVAGGSEGCQATLPICQYFPEKNAEGSLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||:|||||||||||||||||| ||||:|||||||||||:|:||| |:|| || |||
Db         61 QRKTSRNRVYLHTLAESICKLIFPECERLNLALQRTLAKHKIEENRKSSEKEDLEKIIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              :|:||| ||| :|:|||||:|||||||||:|||||||||||||||||||||||||||||
Db        121 EAIAAGAPVEALKDSLGEELFKICYEEDEHILGVVGGTLKDFLNSFSTLLKQSSHCQEAE 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              :|||||||||||||: |||:||||||||||:|:||||||||| :|||: ||||||||||
Db        181 RRGRLEDASILCLDKDQDFLNVYYFFPKRTTALLLPGIIKAAARILYESHVEVSLMPPCF 240

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              :||:|||||||||||:|||||||| |||||||| |||||||||||| |:|: |||
Db        241 RSDCTEFVNQPYLLYSVHVKSTKPSLSPGKPQSSLVIPASLFCKTFPFHFMLDRDLAILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              |||||||:|:|||||||||||:|||||||||||||||||||:|||||||||||||||||
Db        301 LGNGIRRLVNKRDFQGKPNFEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              |||||||||||||||| ||||||||||:|||||||| |||||||||||:|||||||| |
Db        421 AQDGLKKRLGKLKATLEHAHQALEEEKKRTVDLLCSIFPSEVAQQLWQGQIVQAKKFSEV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHR 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLDAY 660

Qy        661 -QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
               || |||| || |||||||||||||||:|
Db        661 HHQGPNSKPWFQDKDVEDGNANFLGKASGVD 691
```

```
RESULT 4
Q9NNW8
ID   Q9NNW8      PRELIMINARY;      PRT;   564 AA.
AC   Q9NNW8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylate cyclase large subunit (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 1-0 FROM N.A.
RX   MEDLINE=20183694; PubMed=10717483;
RA   Zhou Y.H., Zheng J.B., Gu X., Li W.H., Saunders G.F.;
RT   "A novel Pax-6 binding site in rodent B1 repetitive elements:
RT   coevolution between developmental regulation and repeated elements?";
RL   Gene 245:319-328(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Zhou Y.H.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF233750; AAF64043.1; -.
DR   EMBL; AF233746; AAF64043.1; JOINED.
DR   EMBL; AF233747; AAF64043.1; JOINED.
DR   EMBL; AF233748; AAF64043.1; JOINED.
DR   EMBL; AF233749; AAF64043.1; JOINED.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
FT   NON_TER       1      1
SQ   SEQUENCE   564 AA;  63406 MW;  7056E249A75F98D5 CRC64;

  Query Match             82.2%;  Score 2954;  DB 4;  Length 564;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-223;
  Matches  564;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        127 VPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLE 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLE 60

Qy        187 DASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCFHNDCSE 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCFHNDCSE 120

Qy        247 FVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIR 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 FVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIR 180

Qy        307 RLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDL 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDL 240

Qy        367 KGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLK 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLK 300

Qy        427 KRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSD 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSD 360

Qy        487 IVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHA 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHA 420

Qy        547 VQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANK 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 VQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANK 480

Qy        607 FESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAYQQGTNS 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAYQQGTNS 540

Qy        667 KPCFQKKDVEDGNANFLGKASGID 690
              |||||||||||||||||||||||||
Db        541 KPCFQKKDVEDGNANFLGKASGID 564


RESULT 5
Q9PWI2
ID   Q9PWI2      PRELIMINARY;      PRT;   678 AA.
AC   Q9PWI2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase alpha subunit.
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99303623; PubMed=10373466;
RA   Mikami T., Kusakabe T., Suzuki N.;
RT   "Tandem organization of medaka fish soluble guanylyl cylclase alpha1
RT   and beta1 subunit genes. Implications for coordinated transcription of
RT   two subunit genes.";
RL   J. Biol. Chem. 274:18567-18573(1999).
DR   EMBL; AB022280; BAA76690.1; -.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   678 AA;  75166 MW;  66623709CBE68A5C CRC64;

  Query Match             56.1%;  Score 2014.5;  DB 13;  Length 678;
  Best Local Similarity   59.9%;  Pred. No. 2.8e-149;
  Matches  403;  Conservative   99;  Mismatches  152;  Indels   19;  Gaps    8;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||| |||:|||:|||||| |     ||  :    ||::   ::|| :|:  | |:|::|
Db          1 MFCAKLKELKISGECPFSSSAKN---NELGDFEVRSSDAAD-SLPISEDVHGK-IRENVP 55

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQR--TLAKHKIKESRKSL------ERE 112
              ::| ||::| |||| :|| ||  | |:||: |||| | | |       ||
Db         56 RQKASRAKVNLHTLGDSIRKLACPGFQRLHDALQRMMTAADHSKDSESTSLHCCHYQSCS 115

Qy        113 DFEKTIAEQ----AVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFST 168
              |  | : :     :   :|:| :| :||||:| :||||| :|| |||| | ||||||:
Db        116 DKPKHLVDMMKSYSTKTAIPMEALKIALGEEIFSMCYEEDGHILRVVGGALHDFLNSFNV 175

Qy        169 LLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYE 228
              ||||||      :  : : |:|||||:   | ||:| |  || :  ||:|||| :||
Db        176 LLKQSSMLPSPDREDCVNEPSVLCLDKDPGLLTVYFFNPSPTTEMFFPGVIKAAACLLYN 235

Qy        229 TEVEVSLMPPCFHNDCSEFVNQPYLLYSVHMKSTKP-SLSPSKPQSSLVIPTSLFCKTFP 287
              | |:|   |    :   :   || |||:| :|  |  | || :  |:  :|||||   ||
Db        236 TTVDVLKDPLTSKDSILQSSPQPSLLYTVVVKDAKTLSPSPLRATSAGTLPTSLFSTIFP 295

Qy        288 FHFMFDKDMTILQFGNGIRRLMNRRD-FQGKPNFEEYFEILTPKINQTFSGIMTMLNMQF 346
              || : |:|: ::| |:|:|: : |:|  :  | |:|:| |::|:|  || ||:|||| ||
Db        296 FHLILDQDLVLVQIGHGLRRRLTRKDGLRRSPTFQEHFSIVSPQIKCTFQGILTMLNTQF 355

Qy        347 VVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIH 406
              ::|::       : : |||||||||: ||:||||||||||:||: ||||||||||||
Db        356 IIRIKHGVFVTDNTGKPMDLKGQMIYVPESNAILFLGSPCVDKLEELTGRGLYLSDIPIH 415

Qy        407 NALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQL 466
              |||||||:||||:|||||||||||| || || |||||||||||||| ||||  |||||
Db        416 NALRDVVLVGEQAKAQDGLKKRLGKAKAALENAHQALEEEKKKTVDLLFSIFPGTVAQQL 475

Qy        467 WQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVE 526
              |||: ||||||: ||||||||||||||:|| |:|:|||||| |||:|| ||||||||||
Db        476 WQGETVQAKKFNQVTMLFSDIVGFTAVCSLCTPMQVITMLNELYTKFDYQCGELDVYKVE 535

Qy        527 TIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGV 586
              |||||||||||||:||:||||:|| |||||||||||::| ||||:|||||||| |||
Db        536 TIGDAYCVAGGLHRESETHAVEIAFMALKMMELSDEVLTPTGEPIQMRIGLHSGSVLAGV 595

Qy        587 VGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNF 646
              |||:|||||||||||||||||||| |||||:||||:||:|  | ||| ||||::|| ||
Db        596 VGVRMPRYCLFGNNVTLANKFESCSQPRKINISPTTHRLVKGRPEFVFIPRSRQDLPANF 655

Qy        647 PSEIPGICHFLDA 659
              | :|||:|:||::
Db        656 PEDIPGVCYFLES 668


RESULT 6
P79997
ID   P79997      PRELIMINARY;      PRT;   678 AA.
AC   P79997;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase alpha subunit (EC 4.6.1.2).
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98237571; PubMed=9578459;
RA   Mikami T., Kusakabe T., Suzuki N.;
RT   "Molecular cloning of cDNAs and expression of mRNAs encoding alpha and
RT   beta subunits of soluble guanylyl cyclase from medaka fish Oryzias
RT   latipes.";
RL   Eur. J. Biochem. 253:42-48(1998).
DR   EMBL; AB000849; BAA19198.1; -.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   678 AA;  75133 MW;  3A6141D2A6475D40 CRC64;

  Query Match             55.9%;  Score 2009.5;  DB 13;  Length 678;
  Best Local Similarity   59.6%;  Pred. No. 7e-149;
  Matches  401;  Conservative  101;  Mismatches  152;  Indels   19;  Gaps    8;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||| |||:|||:|||||| |     ||  :    ||::   ::|| :|:  | |:|::|
Db          1 MFCAKLKELKISGECPFSSSAKN---NELGDFEVRSSDAAD-SLPISEDVHGK-IRENVP 55

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQR--TLAKHKIKESRKSL------ERE 112
              ::| ||::| |||| :|| ||  | |:||: |||| | | |       ||
Db         56 RQKASRAKVNLHTLGDSIRKLACPGFQRLHDALQRMMTAADHSKDSESTSLHCCHYQSCS 115

Qy        113 DFEKTIAEQ----AVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFST 168
              |  | : :     :   :|:| :| :||||:| :||||| :|| |||| | ||||||:
Db        116 DKPKHLVDMMKSYSTKTAIPMEALKIALGEEIFSMCYEEDGHILRVVGGALHDFLNSFNV 175

Qy        169 LLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYE 228
              ||||||      :   : : |:|||||:   | ||:| |  || :  ||:||||| :||
```

```
Db        176 LLKQSSMLPSPDREDCVNEPSVLCLDKDPGLLTVYFFNPSPTTEMFFPGVIKAAACLLYN 235

Qy        229 TEVEVSLMPPCFHNDCSEFVNQPYLLYSVHMKSTKP-SLSPSKPQSSLVIPTSLFCKTFP 287

Db        236 TTVDVLKDPLTSKDSILQSSPQPSLLYTVVVKDAKTLSPSPLRATSAGTLPTSLFSTIFP 295

Qy        288 FHFMFDKDMTILQFGNGIRRLMNRRD-FQGKPNFEEYFEILTPKINQTFSGIMTMLNMQF 346

Db        296 FHLILDQDLVLVQIGHGLRKRLTRKDGLRRSPTFQEHFSIVSPQIKCTFQGILTMLNTQF 355

Qy        347 VVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIH 406

Db        356 IIRIKHGVFVTDNTGKPMDLKGQMIYVPESNAILFLGSPCVDKLEELTGRGLYLSDIPIH 415

Qy        407 NALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQL 466

Db        416 NALRDVVLVGEQAKAQDGLKKRLGKAKAALENAHQALEEEKKKTVDLLFSIFPGTVAQQL 475

Qy        467 WQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVE 526

Db        476 WQGETVQAKKFNQVTMLFSDIVGFTAVCSLCTPMQVITMLNELYTKFDYQCGELDVYKVE 535

Qy        527 TIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGV 586

Db        536 TIGDAYCVAGGLHRESETHAVEIAFMALKMIELSDEVLTPTGEPIQMRIGLHSGSVLAGV 595

Qy        587 VGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNF 646

Db        596 VGVRMPRYCLFGNNVTLANKFESCSQPRKVNISPTTHRLVKGRPEFVFIPRSRQDLPANF 655

Qy        647 PSEIPGICHFLDA 659

Db        656 PEDIPGVCYFLES 668


RESULT 7
Q90VV5
ID   Q90VV5      PRELIMINARY;      PRT;   675 AA.
AC   Q90VV5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Soluble guanylyl cyclase alpha1 subunit.
GN   FRGCS-ALPHA1.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morinaga C., Yamamoto T., Moriya Y., Suzuki N.;
RT   "Identification of tandem organization of soluble guanylyl cyclase
RT   alpha1 and beta1 subunit genes in the Japanese pufferfish (Fugu
RT   rubripes) genome.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB062171; BAB60907.1; -.
DR   EMBL; AB062169; BAB60905.1; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   675 AA;  75498 MW;  E71A283DC0369601 CRC64;

  Query Match           55.8%;  Score 2005;  DB 13;  Length 675;
  Best Local Similarity   60.1%;  Pred. No. 1.6e-148;
  Matches  406;  Conservative   96;  Mismatches  147;  Indels   26;  Gaps    9;

Qy          1 MFCTKLKDLKITGECPFSLLA-PGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESL 59

Db          1 MFCTKLKELKISGECPFTSSAKTNQL--EHSEQLSKD----VAHLPTSKDV-RRNTGENA 53

Qy         60 PQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHK-----------IKESRK 107

Db         54 PHQKSSRAKVNLHTLGESIRKLACPEFERLHTALRRMMRLSDPSREPASPFVCCTEYQRC 113

Qy        108 SLEREDFEKTIAEQAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFS 167

Db        114 SDDPEHFVKMMSMFSSRKAIQMDTLRVALGEELFNMCYEEDRHILRVVGGALHDFLNSFN 173

Qy        168 TLLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLY 227

Db        174 VLLKQSSTLHYQDREDCVNEPSVLCLDKDLGLLTVYFFNPHPTTELFFPGVIKAAARLLY 233

Qy        228 ETEVEVSLMPPCFHNDCSEFVNQPYLLYSVHMKSTKPSLSPS--KPQSSLVIPTSLFCKT 285

Db        234 HTTVDVLMDPPATKDSILQSSPQPSLLYTVVVKDAK-SLSPSPLRATSAGTLPTSLFSTI 292

Qy        286 FPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKP-NFEEYFEILTPKINQTFSGIMTMLNM 344

Db        293 FPFHLILDQDLVLVQVGHGLRKRLTRKDGLRRPATFQEHFSILSPLIRCTFQGILTMLNT 352

Qy        345 QFVVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIP 404

Db        353 QFTIRIKRGVSTA--DNTLMDLKGQMIYVPESDAILFLGSPCVDKLEELTGRGLYLSDIP 410

Qy        405 IHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQ 464

Db        411 IHNALRDVVLVGEQAKAQDGLKKRLGKAKAALEHAHQALEEEKKKTVDLLFTIFPGTVAQ 470

Qy        465 QLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYK 524

Db        471 QLWQGQTVQAKKFERVTMLFSDIVGFTAVCSHCTPMQVITMLNELYTKFDHHCGELDVYK 530

Qy        525 VETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFA 584

Db        531 VETIGDAYCVAGGLHKESETHAVQVALMALKMMELSNGVMTPAGEPIQMRIGLHTGSVLA 590

Qy        585 GVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPP 644

Db        591 GVVGVKMPRYCLFGNNVTLANKFESCSQPGKINISPTTHRLLEDHPEFVFIPRRRQELPA 650

Qy        645 NFPSEIPGICHFLDA 659

Db        651 NFPEDIPGVCYFLEA 665


RESULT 8
O77105
ID   O77105      PRELIMINARY;      PRT;   699 AA.
AC   O77105;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase alpha-1 subunit.
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingiodea; Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98409773; PubMed=9736646;
RA   Nighorn A., Gibson N.J., Rivers D.M., Hildebrand J.G., Morton D.B.;
RT   "The nitric oxide-cGMP pathway may mediate communication between
RT   sensory afferents and projection neurons in the antennal lobe of
RT   manduca sexta.";
RL   J. Neurosci. 18:7244-7255(1998).
DR   EMBL; AF062750; AAC61263.1; -.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
```

```
KW   Lyase.
SQ   SEQUENCE   699 AA;  78490 MW;  989F8067ABAF8465 CRC64;

  Query Match             30.5%;  Score 1097;  DB 5;  Length 699;
  Best Local Similarity   37.6%;  Pred. No. 2.5e-77;
  Matches  266;  Conservative  130;  Mismatches  262;  Indels   50;  Gaps   19;

Qy         15 CPF------SLLAPGQVPNESSEEAAGSSESCKATVPICQD-----IPEKNIQESLPQRK 63
              |||      | |       |    : |   | |  :|     :  |:: |:| |
Db          3 CPFRRASSQHQFANGGSSAPKKPEFRSRTSSVHLTGPEEEDGERNTLTLKHMSEAL-QLL 61

Qy         64 TSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAEQAV 123
              |: |   ||    |: |     : : |          :: :  |:     : : | :
Db         62 TAPSNECLHAAVTSLTKNQSDHYHKYN-------CLRRLPDDVKTCRNYAYLQEIYDAVR 114

Qy        124 AA-GVPVEVIKESLGEEVFKICYEED---ENILGVVGGTLKDFLNSFSTLLKQSSHCQEA 179
              |   | :     ||| :    :  :   |     :|  | :|| :  : :    | |:
Db        115 ATDSVNTKDFMAKLGEYLILTAFSHNCRLERAFKCLGTNLTEFLTTLDS-VHDVLHDQDT 173

Qy        180 G-KRGRLE-DASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMP 237
              |   :| :|: :|   ::  : ::        : :| | :|| |  ||:|: :: |
Db        174 PLKDETMEYEANFVCTTSQEGKIQLHLTTESEPVAYLLVGSLKAIAKRLYDTQTDIRLR- 232

Qy        238 PCFHNDCSEFVNQPYLLYSV--HMKSTK---------PSLSPSKPQSSLVIPTSLFCKTF 286
               : ||   |    | : :|  | || :          |:: |   :  | |  : ||| |
Db        233 -SYTNDPRRF---RYEINAVPLHQKSKEDSCELVNEAASVATSTKVTDLKIGVASFCKAF 288

Qy        287 PFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPK-INQTFSGIMTMLNMQ 345
              |:||: || : ::| | |  ||          :   || :| |: :   |  |: :|
Db        289 PWHFITDKRLELVQLGAGFMRLFGTHLATHGSSLGTYFRLLRPRGVPLDFREILKRVNTP 348

Qy        346 FVVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPI 405
              |:  ::   ::       :::||||::  || ::||:||| :| ||  |||||::||||:
Db        349 FMFCLKMPGSTALAEG--LEIKGQMVFCAESDSLLFVGSPFLDGLEGLTGRGLFISDIPL 406

Qy        406 HNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQ 465
              |:| |||:|:||||||||||::|: ||| ::|:| :|:::|::| | ||  |||  :|::
Db        407 HDATRDVILVGEQARAQDGLRRRMDKLKNSIEEASKAVDKEREKNVSLLHLIFPPHIAKR 466

Qy        466 LWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKV 525
              || |: ::||   :|||||||||||:||:  :|: || ||  ||: ||  | |||||||
Db        467 LWLGEKIEAKSHDDVTMLFSDIVGFTSICATATPMMVIAMLEDLYSVFDIFCEELDVYKV 526

Qy        526 ETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAG 585
              |||||||||| |||:: :||| ||| |||:|:|   : ::  | |||||||||:|:| ||
Db        527 ETIGDAYCVASGLHRKVETHAPQIAWMALRMVETCAQHLTHEGNPIKMRIGLHTGTVLAG 586

Qy        586 VVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPN 645
              |||  | :|||||:|||||||||| | | |||||||||  |   |||  || |  || :
Db        587 VVGKTMLKYCLFGHNVTLANKFESGSEPLKINVSPTTYEWLIKFPGFDMEPRDRSCLPNS 646

Qy        646 FPSEIPGICHFLDAY-QQGTN-SKPCFQKKDVEDGNANF-LGKASGID 690
              || :| | |:||  |   ||:  :|  | | : :   :: :|:|:  |
Db        647 FPKDIHGTCYFLHKYTHPGTDPGEP--QVKHIREALKDYGIGQANSTD 692


RESULT 9
Q24085
ID   Q24085      PRELIMINARY;      PRT;   676 AA.
AC   Q24085;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   GYCALPHA99B protein.
GN   GYC-ALPHA-99B OR GYCALPHA99B OR DGCA1 OR CG1912.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=95318108; PubMed=7797526;
RA   Shah S., Hyde D.R.;
RT   "Two Drosophila genes that encode the alph and beta subunits of the
RT   brain soluble guanylyl cyclase.";
RL   J. Biol. Chem. 270:15368-15376(1995).
DR   EMBL; AE003770; AAF56917.1; -.
DR   EMBL; U27117; AAA87940.1; -.
DR   HSSP; P19687; 1AWN.
DR   FlyBase; FBgn0013972; Gyc-alpha-99B.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   676 AA;  75662 MW;  CE8097E1EC3787F8 CRC64;

  Query Match             27.6%;  Score 991.5;  DB 5;  Length 676;
  Best Local Similarity   37.8%;  Pred. No. 4.4e-69;
  Matches  247;  Conservative  117;  Mismatches  209;  Indels   81;  Gaps   20;

Qy         58 SLPQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAK--------HKIK---ESR 106
              :|   : |   : |  | :|  |  |  | || |:   :||        ||:|   ::
Db         27 ALEDEELSDDALTLTHLQMAIQLLTAPSNEDLNTAVTSLVAKYRQNWPNIHKLKLDPQTF 86

Qy        107 KSLEREDFEKTIAE------QAVAAGVPVEVIKESLGEEVFKICYEE-DENILGVVGGTL 159
              ||    |:   | |       :| |: : |      ||||:   |     |     :|  |
Db         87 KSCANYDYLADIQELLLKMDEASASEILV-----LLGEELITCCCTGIIERAFRCLGTDL 141

Qy        160 KDFLNSFS---TLLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILP 216
              ::|| |      :||      ||     : |   :|  : :    :  :   :  : :|
Db        142 QEFLGSLDGVYDVLK----LQEED----VTDTGFVCAGEGE----LIFTSERPVIAWLLL 189
```

```
Qy        217 GIIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQPYLLYSVH-----MKSTKP-SLSPSK 270
Db        190 GSLKALTRMLYKVDVNIKIEP--VEGDARRY---RYLFSLVKDNSQTMLMGRPTSVSKTI 244

Qy        271 PQ----------SSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
Db        245 PETVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLY-------KPYM 297

Qy        321 EE-------YFEILTPK-INQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIY 372
Db        298 ADFGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIGLEIKGQMVH 357

Qy        373 IVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKL 432
Db        358 CPESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQARAQDGLRRRMDKI 417

Qy        433 KATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTA 492
Db        418 KNSIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYPDVTILFSDIVGFTS 477

Qy        493 ICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALM 552
Db        478 ICSRATPFMVISMLEGLYKDFDEFCDFFDVYKVETIGDAYCVASGLHRASIYDAHKVAWM 537

Qy        553 ALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
Db        538 ALKMIDACSKHITHDGEQIKMRIGLHTGTVLAGVVGRKMPRYCLFGHSVTIANKFESGSE 597

Qy        613 PRKINVSPTTYRLLKDCPGFVF--TPRSREELPPNFPSEIPG---ICHFLDAYQ 661
Db        598 ALKINVSPTTKDWLTKHEGFEFELQPRDPSFLPKEFPN--PGGTETCYFLESFR 649
```

RESULT 10
Q95SQ4
ID Q95SQ4 PRELIMINARY; PRT; 676 AA.
AC Q95SQ4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GH08311p.
GN GYC-ALPHA-99B OR GYCALPHA99B OR CG1912.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY060654; AAL28202.1; -.
DR FlyBase; FBgn0013972; Gyc-alpha-99B.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ SEQUENCE 676 AA; 75658 MW; CB9183E1EC314138 CRC64;

Query Match 27.4%; Score 985.5; DB 5; Length 676;
Best Local Similarity 37.6%; Pred. No. 1.3e-68;
Matches 246; Conservative 117; Mismatches 210; Indels 81; Gaps 20;

```
Qy         58 SLPQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAK--------HKIK---ESR 106
Db         27 ALEDEELSDDALTLTHLQMAIQLLTAPSNEDLNTAVTSLVAKYRQNWPNIHKLKLDPQTF 86

Qy        107 KSLEREDFEKTIAE------QAVAAGVPVEVIKESLGEEVFKICYEE-DENILGVVGGTL 159
```

```
Db         87 KSCANYDYLADIQELLLKMDEASASEILV-----LLGEELITCCCTGIIERAFRCLGTDL 141

Qy        160 KDFLNSFS---TLLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILP 216
Db        142 QEFLGSLDGVYDVLK----LQEED----VTDTGFVCAGEGE----LIFTSERPVIAWLLL 189

Qy        217 GIIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQPYLLYSVH-----MKSTKP-SLSPSK 270
Db        190 GSLKALTRMLYKVDVNIKIEP--VEGDARRY---RYLFSLVKDNSQTMLMGRPTSVSKTI 244

Qy        271 PQ----------SSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
Db        245 PETVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLY-------KPYM 297

Qy        321 EE-------YFEILTPK-INQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIY 372
Db        298 ADFGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIGLEIKGQMVH 357

Qy        373 IVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKL 432
Db        358 CPESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQARAQDGLRRRMDKI 417

Qy        433 KATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTA 492
Db        418 KNSIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYPDVTILFSDIVGFTS 477

Qy        493 ICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALM 552
Db        478 ICSRATPFMVISMLEGLYKDFDEFCDFFDVYKVETIGDAYCVASGLHRASIYDAHKVAWM 537

Qy        553 ALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
Db        538 ALKMIDACSKHITHDGEQIKMRIGLHTGTVLAGVVGRKMPRYCLFGHSVTIANKFESGSE 597

Qy        613 PRKINVSPTTYRLLKDCPGFVF--TPRSREELPPNFPSEIPG---ICHFLDAYQ 661
Db        598 ALKINVSPTPKDWLTKHEGFEFELQPRDPSFLPKEFPN--PGGTETCYFLESFR 649
```

RESULT 11
Q8R5L4
ID Q8R5L4 PRELIMINARY; PRT; 220 AA.
AC Q8R5L4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Soluble guanylate cyclase alpha2d (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA Okamoto H., Asakawa T.;
RT "Molecular cloning of a novel variant of the rat soluble guanylate
RT cyclase beta2 subunit.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB079780; BAB84824.1; -.
FT NON_TER 1 1
FT NON_TER 220 220
SQ SEQUENCE 220 AA; 24646 MW; D8BEC749FDA5E35E CRC64;

Query Match 22.9%; Score 823; DB 11; Length 220;
Best Local Similarity 60.7%; Pred. No. 1.5e-56;
Matches 162; Conservative 30; Mismatches 27; Indels 48; Gaps 1;

```
Qy        348 VRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHN 407
Db          1 IRTKPEASGTDNEDKVMEIKGQMIHVPESNAILFLGSPCVDKLDELIGRGLHLSDIPIHD 60

Qy        408 ALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLW 467
```

```
Db         61 ATRDVILVGEQAKAQDGLKKRMDKLKATLEKTHQALEEEKKKTVDLLYSIFPGDVAQQLW 120

Qy        468 QGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVET 527

Db        121 QRQQVQARKFDDVTMLFSDIVGFTAICAQCTPMQVISMLNELYTRFDHQCGFLDIY---- 176

Qy        528 IGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVV 587

Db        177 ----------------------------------------KMRIGIHSGSVLAGVV 192

Qy        588 GVKMPRYCLFGNNVTLANKFESCSVPR 614

Db        193 GVRMPRYCLFGNNVTLASKFESGSHPR 219


RESULT 12
Q90VY5
ID   Q90VY5      PRELIMINARY;      PRT;   617 AA.
AC   Q90VY5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta1 subunit.
GN   FRGCS-BETA1.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morinaga C., Yamamoto T., Moriya Y., Suzuki N.;
RT   "Identification of tandem organization of soluble guanylyl cyclase
RT   alpha1 and beta1 subunit genes in the Japanese pufferfish (Fugu
RT   rubripes) genome.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB062172; BAB60908.1; -.
DR   EMBL; AB062170; BAB60906.1; -.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR000834; Zn_carbOpept.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   PROSITE; PS00132; CARBOXYPEPT_ZN_1; UNKNOWN_1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   617 AA;  70236 MW;  7B531B5896A06191 CRC64;

  Query Match           21.9%;  Score 788;  DB 13;  Length 617;
  Best Local Similarity   34.2%;  Pred. No. 3.5e-53;
  Matches  202;  Conservative  102;  Mismatches  217;  Indels   70;  Gaps   15;

Qy        103 KESRKSLEREDFEKTIAEQA-----VAAG-----VPVEVIKESLGEEVFKICYEED-ENI 151

Db         27 REAQLDIEGQFLVRIIYEDAKTYDLVAAASKVLKIDAGDILQLFGKMFFEFCQESGYDTI 86

Qy        152 LGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRT 210

Db         87 LRVLGSNVREFLQNLDAL-----HDHLGTIYPGMRAPSFRCTDAEKGNSLILHYYSEREG 141

Qy        211 TSLILPGIIKAAAHVLYETEVEVSLMPP----CFHNDCSEFVNQPYLLYSVHMKSTK--- 263

Db        142 LQDIVIGIIKTVAQQIHGTEIEMKMIQPKSKECDH----------IKFLIEEKDSEEEA 190

Qy        264 -----PSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKP 318

Db        191 FYEDLDGFEENGTQETRISPYT-FCKAFPFHLMFDRDLMLTQCGNAIYRVLPQLQ-PGSC 248

Qy        319 NFEEYFEILTPKINQTFSGIMTMLNMQFVVR----------VRRWDNSVKKSSRVMDLKG 368

Db        249 ILPSVFSLVRPHIDFSFHGILSHINTVFVLRSKEGLLNVETVENEDELTGVEISCLRLKG 308

Qy        369 QMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKR 428

Db        309 QMIYLPEAENILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQE 368

Qy        429 LGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIV 488

Db        369 LEILTDRLQHTLRALEDEKKKTDRLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIV 428

Qy        489 GFTAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKE 541

Db        429 GFNTFCSKHASAEGAIKIVNLLNDVYTRFDILTDSRNNPYVYKVETVGDKYMTVSGLPEP 488

Qy        542 SDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNV 601

Db        489 CTHHAKSICHLALEMLEIAGQV-KVDDKPVQITIGIHTGEVVTGVIGQRMPRYCLFGNTV 547

Qy        602 TLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652

Db        548 NLTSRTETTGEKGRINVSEFTFRCL----------QSAENADPQFHLEYRG 588


RESULT 13
O77106
ID   O77106      PRELIMINARY;      PRT;   600 AA.
AC   O77106;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta-1 subunit.
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingiodea; Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98409773; PubMed=9736646;
RA   Nighorn A., Gibson N.J., Rivers D.M., Hildebrand J.G., Morton D.B.;
RT   "The nitric oxide-cGMP pathway may mediate communication between
RT   sensory afferents and projection neurons in the antennal lobe of
RT   manduca sexta.";
RL   J. Neurosci. 18:7244-7255(1998).
DR   EMBL; AF062751; AAC61264.1; -.
DR   HSSP; P16068; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   600 AA;  68099 MW;  7F7D0136E6ADACB9 CRC64;

  Query Match           21.9%;  Score 785.5;  DB 5;  Length 600;
  Best Local Similarity   33.9%;  Pred. No. 5.2e-53;
  Matches  196;  Conservative  106;  Mismatches  217;  Indels   59;  Gaps   13;

Qy         89 LNVALQRTLAKHKIKESRKSLE------------REDFEKTIAEQAVAAGV-----PVE 130

Db          5 VNYALELLVMKTFDEETWETIKKKADVAMEGSFLVRQIYEDEITYNLITAAVEVLQIPAD 64

Qy        131 VIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDAS 189

Db         65 AILELFGKTFFEFCQDSGYDKILQVLGATPRDFLQNLDGL-----HDHLGTLYPGMRSPS 119

Qy        190 ILCLDK-EDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCFHNDCSEF- 247

Db        120 FRCTERPEDGALVLHYYSDRPGLEHIVIGIVKTVASKLHNTEVKVEILKTKEECDHVQFL 179

Qy        248 ---------VNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTI 298

Db        180 ITETSTTGRVSAPEIA-EIETLSLEPKVSP-----------ATFCRVFPFHLMFDRDLNI 227

Qy        299 LQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSV- 357

Db        228 VQAGRTVSRLLPRVTRPG-CKITDVLDTVRPHLEMTFANVLAHINTVYVLKTKPEEMSVT 286
```

```
Qy        358 --KKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLI 415
                  :    : |||||:|| |:  ::|   | |  |:| | ||| ::|||:|:| ||:||:
Db        287 DPHEEIASLRLKGQMLYIPETDVVVFQCYPSVTNLDDLTRRGLCIADIPLHDATRDLVLM 346

Qy        416 GEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAK 475
              ||  |   | : |  |   |:|  : || ||:||  || |: |  || :|   : | |:
Db        347 SEQFEADYKLTQNLEVLTDKLQQTFRELELEKQKTDRLLYSVLPISVATELRHRRPVPAR 406

Qy        476 KFSNVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFD---QQCGELDVYKVETI 528
              ::  ||:||| |||   |:: |     :::: ||| ||| ||          :||||||:
Db        407 RYDTVTLLFSGIVGFANYCARNSDHKGAMKIVRMLNDLYTAFDVLTDPKRNPNVYKVETV 466

Qy        529 GDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVG 588
              || |    || :    ||  |:|:|| ||:|| : ::  |||: : ||:||| | . ||:|
Db        467 GDKYMAVSGLPEYEVAHAKHISLLALDMMDLS-QTVTVDGEPVGITIGIHSGEVVTGVIG 525

Qy        589 VKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLL 626
              :|||||||| | | :: |:  ||  ||||  ||  |
Db        526 HRMPRYCLFGNTVNLTSRCETTGVPGTINVSEDTYNYL 563


RESULT 14
Q91XJ7
ID   Q91XJ7      PRELIMINARY;      PRT;   742 AA.
AC   Q91XJ7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta 2 subunit.
GN   GUCY1B2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX   MEDLINE=21391941; PubMed=11406623;
RA   Koglin M., Vehse K., Budaeus L., Scholz H., Behrends S.;
RT   "Nitric Oxide Activates the beta 2 Subunit of Soluble Guanylyl Cyclase
RT   in the Absence of a Second Subunit.";
RL   J. Biol. Chem. 276:30737-30743(2001).
DR   EMBL; AY004153; AAF86581.1; -.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ   SEQUENCE   742 AA;  83050 MW;  41CE148B15E244E4 CRC64;

  Query Match             21.8%;  Score 784.5;  DB 11;  Length 742;
  Best Local Similarity   32.6%;  Pred. No. 8.5e-53;
  Matches  194;  Conservative  107;  Mismatches  213;  Indels   81;  Gaps   10;

Qy        116 KTIAEQAVAAGVPVEVIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSS 174
              | | |      | :| | :  ||  || |     : :|  :|| | :|: :   |
Db         48 KLIQEACKVLDVSMEAILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDAL----- 102

Qy        175 HCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVS 234
              |   |      :   |    :  | : ::|:  :      |:||||:| |   ::|:| :|
Db        103 HSYLALSYQEMNAPSFRVEEGADGAMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMS 162

Qy        235 LMPPCFHNDCSEFV----NQPYLLYSVHMKSTKP--SLSPSKPQSS-------------- 274
              ::      | :| |     : :::: |  |: :       |:|| |
Db        163 IL------DMNEEVERTGKKEHVVFLVVQKAHRQIRGAKASRPQGSEDSQADQEALQGTL 216

Qy        275 ---------------------------------------------LVIPTSLFCKTFPFH 289
                                                          | :  :|| ||||
Db        217 LRMKERYLNIPVCPGEKSHSTAVRASVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFH 276

Qy        290 FMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVR 349
```

```
                :||: : : | |  |:: :     | |   :||| |: |::    | |   :| |||::
Db        277 IVFDEALRVKQAGVNIQKYVPGILTQ-KFALDEYFSIIHPQVTFNISSICKFINSQFVLK 335

Qy        350 VRR-WDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNA 408
              |:       :||  :: |:||||::     ::|: || |  |::     ::|||  |:
Db        336 TRKEMMPKARKSQPMLKLRGQMIWMESLRCMIFMCSPNVRSLQELEESKMHLSDIAPHDT 395

Qy        409 LRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQ 468
              ||::|: :|  |:  |  :| | |  |        |  |||||  || :: |  || || :
Db        396 TRDLILLNQQRLAEMELSCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVANQLKE 455

Qy        469 GQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETI 528
              |: | | :|   |:||||:| || ||: | |:|:: |||::|::||:     ||||||||
Db        456 GRKVAAGEFETCTILFSDVVTFTNICAACEPIQIVNMLNSMYSKFDRLTSVHDVYKVETI 515

Qy        529 GDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSP-HGEPIKMRIGLHSGSVFAGVV 587
              |||| | ||:    ::|| ::|  || |   : |||:|  ||||::|:|:|:| | ||||
Db        516 GDAYMVVGGVPVPVESHAQRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVV 575

Qy        588 GVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              | |||||||||: |  |:: ||  :| |:::||| :| ||:  ||    |  |:
Db        576 GDKMPRYCLFGDTVNTASRMESHGLPSKVHLSPTAHRALKN-KGFEIVRRGEIEV 629


RESULT 15
O54865
ID   O54865      PRELIMINARY;      PRT;   620 AA.
AC   O54865;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylate cyclase beta-1 subunit.
GN   GUCY1B3 OR GC-S-BETA-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Gansemans Y., Brouckaert P., Fiers W.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/10;
RX   PubMed=10984516;
RA   Sharina I.G., Krumenacker J.S., Martin E., Murad F.;
RT   "Genomic organization of alpha 1 and beta 1 subunits of the mammalian
RT   soluble guanylyl cyclase genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10878-10883(2000).
DR   EMBL; AF020339; AAB94876.1; -.
DR   EMBL; AF297083; AAG17447.1; -.
DR   HSSP; P16068; 1AWN.
DR   MGD; MGI:1860604; Gucy1b3.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   620 AA;  70597 MW;  EED89ABF539F4EF9 CRC64;

  Query Match             21.8%;  Score 781.5;  DB 11;  Length 620;
  Best Local Similarity   35.5%;  Pred. No. 1.1e-52;
  Matches  200;  Conservative   97;  Mismatches  209;  Indels   57;  Gaps   14;

Qy        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153
              . :: |  : | |  :: :     |||   |      |:: :  |:  |  | |   : ||
Db         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88

Qy        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212
              |:|  :::|| :   |      |   |      :   | | | |    | ::|:  :
Db         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143
```

```
Qy        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265
Db        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192

Qy        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
Db        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRNLVVTQCGNAIYRVLPQLQ-PGNCSL 250

Qy        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370
Db        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGAEISCLRLKGQM 310

Qy        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430
Db        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370

Qy        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490
Db        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430

Qy        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543
Db        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490

Qy        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603
Db        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549

Qy        604 ANKFESCSVPRKINVSPTTYRLL 626
Db        550 TSRTETTGEKGKINVSEYTYRCL 572
```

Search completed: June 27, 2003, 13:04:41
Job time : 38.7899 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:32 ; Search time 14.2322 Seconds
(without alignments)
1426.467 Million cell updates/sec

Title: US-09-762-767A-2
Perfect score: 3593
Sequence: 1 MFCTKLKDLKITGECPFSLL.....QKKDVEDGNANFLGKASGID 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	407.5	11.3	1075	1	US-07-623-033-2	Sequence 2, Appli
2	325.5	9.1	1077	4	US-09-412-210-1	Sequence 1, Appli
3	322	9.0	1064	3	US-08-726-214-8	Sequence 8, Appli
4	319.5	8.9	1099	3	US-08-726-214-14	Sequence 14, Appl
5	319	8.9	1090	3	US-08-307-896-3	Sequence 3, Appli
6	319	8.9	1090	3	US-08-726-214-4	Sequence 4, Appli
7	319	8.9	1090	5	PCT-US95-11808-3	Sequence 3, Appli
8	303	8.4	1144	3	US-08-726-214-6	Sequence 6, Appli
9	302	8.4	1134	3	US-08-726-214-2	Sequence 2, Appli
10	299.5	8.3	533	3	US-08-726-214-18	Sequence 18, Appl
11	287.5	8.0	1253	4	US-08-864-785-2	Sequence 2, Appli
12	284.5	7.9	1248	3	US-08-726-214-16	Sequence 16, Appl
13	283.5	7.9	1098	3	US-08-726-214-10	Sequence 10, Appl
14	282.5	7.9	604	4	US-09-008-097-4	Sequence 4, Appli
15	282.5	7.9	1167	4	US-09-008-097-6	Sequence 6, Appli
16	280.5	7.8	1261	4	US-09-473-716-2	Sequence 2, Appli
17	279.5	7.8	1168	4	US-09-474-076-2	Sequence 2, Appli
18	279.5	7.8	1180	3	US-08-726-214-12	Sequence 12, Appl
19	272.5	7.6	1165	1	US-08-240-357-2	Sequence 2, Appli
20	263.5	7.3	1294	4	US-09-473-717-2	Sequence 2, Appli
21	260.5	7.3	1305	4	US-08-864-785-3	Sequence 3, Appli
22	260.5	7.3	1353	3	US-08-894-173-2	Sequence 2, Appli
23	260.5	7.3	1353	4	US-09-398-193-2	Sequence 2, Appli
24	260.5	7.3	1353	4	US-09-473-717-3	Sequence 3, Appli
25	255.5	7.1	1353	4	US-09-398-193-99	Sequence 99, Appl
26	195	5.4	670	2	US-08-997-080-178	Sequence 178, App
27	195	5.4	670	2	US-08-997-362-178	Sequence 178, App
28	195	5.4	670	4	US-09-095-855-178	Sequence 178, App
29	195	5.4	670	4	US-09-324-542-178	Sequence 178, App
30	195	5.4	670	4	US-09-205-426-178	Sequence 178, App
31	195	5.4	722	2	US-08-997-080-174	Sequence 174, App
32	195	5.4	722	2	US-08-997-362-174	Sequence 174, App
33	195	5.4	722	4	US-09-095-855-174	Sequence 174, App
34	195	5.4	722	4	US-09-324-542-174	Sequence 174, App
35	195	5.4	722	4	US-09-205-426-174	Sequence 174, App
36	192.5	5.4	107	4	US-08-857-076-76	Sequence 76, Appl
37	191	5.3	112	4	US-08-857-076-75	Sequence 75, Appl
38	187	5.2	419	2	US-08-997-080-125	Sequence 125, App
39	187	5.2	419	2	US-08-997-362-125	Sequence 125, App
40	187	5.2	419	4	US-09-095-855-125	Sequence 125, App
41	187	5.2	419	4	US-09-324-542-125	Sequence 125, App
42	187	5.2	419	4	US-09-205-426-125	Sequence 125, App
43	180	5.0	429	3	US-08-307-896-4	Sequence 4, Appli
44	180	5.0	429	5	PCT-US95-11808-4	Sequence 4, Appli
45	138.5	3.9	85	3	US-08-894-173-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-07-623-033-2
; Sequence 2, Application US/07623033
; Patent No. 5237051
; GENERAL INFORMATION:
; APPLICANT: Garbers, David L.
; APPLICANT: Schulz, Stephanie
; TITLE OF INVENTION: CLONING THE ENTEROTOXIN RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGMUS & CHESTNUT
; STREET: 100 South Wacker Drive - Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/623,033
; FILING DATE: 19901206
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: VU9018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 456-8000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1075 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-623-033-2

Query Match 11.3%; Score 407.5; DB 1; Length 1075;
Best Local Similarity 36.6%; Pred. No. 3.1e-31;
Matches 98; Conservative 47; Mismatches 94; Indels 29; Gaps 6;

```
Qy        423 DGLKKRL----GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFS 478
              | | :||      |:  :|:  |  : |: :   |   : |  | : | :  :|: : :
Db        764 DTLIRRLQLYSRNLEHLVEERTQLYKAERDRADHLNFMLLPRLVVKSLKEKGIVEPELYE 823

Qy        479 NVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL 538
                ||: |||||||| ||   :|::|: ||| :|  |||     ||||||||||| || ||
Db        824 EVTIYFSDIVGFTTICKYSTPMEVVDMLNDIYKSFDQIVDHHDVYKVETIGDAYVVASGL 883
```

```
Qy        539 -HKESDTHAVQIALMALKMMELSDEVMSPH--GEPIKMRIGLHSGSVFAGVVGVKMPRYC 595
              :  : ||| |: ||| ::         |  | |: :|||:|||   |||||:||||||
Db        884 PMRNGNRHAVDISKMALDILSFMGTFELEHLPGLPVWIRIGVHSGPCAAGVVGIKMPRYC 943

Qy        596 LFGNNVTLANKFESCSVPRKINVSPTTYRLLK--DCPGFVFTPRSREELPPNFPSEIPGI 653
              |||: |  |:: ||  :| :|::| :|  :|:  ||                |  |: |
Db        944 LFGDTVNTASRMESTGLPLRIHMSSSTIAILRRTDC--------------QFLYEVRG- 987

Qy        654 CHFLDAYQQGTNSKPCFQKKDVEDGNAN 681
                : | :|  ::  :    ::|   |
Db        988 ----ETYLKGRGTETTYWLTGMKDQEYN 1011


RESULT 2
US-09-412-210-1
; Sequence 1, Application US/09412210
; Patent No. 6403358
; GENERAL INFORMATION:
;  APPLICANT:  Kapeller-Libermann, Rosana
;  TITLE OF INVENTION:  21529, A NOVEL ADENYLATE CYCLASE
;  FILE REFERENCE:  5800-47
;  CURRENT APPLICATION NUMBER: US/09/412,210
;  CURRENT FILING DATE:  1999-10-05
;  NUMBER OF SEQ ID NOS:  2
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1077
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-412-210-1

  Query Match             9.1%;  Score 325.5;  DB 4;  Length 1077;
  Best Local Similarity   28.2%;  Pred. No. 4.6e-23;
  Matches   91;  Conservative   55;  Mismatches   98;  Indels   79;  Gaps    9;

Qy        365 DLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDG 424
              | :  ::  : ::|:|||             |:|      | || :        ||
Db        165 DSRPALLPQLAANAVLFLCGNVA---------GVY------HKALME--------RA--- 198

Qy        425 LKKRLGKLKATLEQA------HQALEEEKKKTVDLLCSIFP--------CEVAQQLWQGQ 470
                    |:||  :|       : |: |||    || || |          |:  :|  ||
Db        199 -------LRATFREALSSLHSRRRLDTEKKHQEHLLLSILPAYLAREMKAEIMARLQAGQ 251

Qy        471 -----------VVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGE 519
                        :  |:   |::|::|||||| : |:||| ::: ||| |: :|||   |
Db        252 GSRPESTNNFHSLYVKRHQGVSVLYADIVGFTRLASECSPKELVLMLNELFGKFDQIAKE 311

Qy        520 LDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHS 579
              :  ::: :|| |    ||      ||:    | | |     :: :  |  | ||:|:||
Db        312 HECMRIKILGDCYYCVSGLPLSLPDHAINCVRMGLDMCRAIRKLRAATGVDINMRVGVHS 371

Qy        580 GSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDC---------- 629
              |||  ||:|::  :| :: ::|||||  |:  || :::::  |  ||
Db        372 GSVLCGVIGLQKWQYDVWSHDVTLANHMEAGGVPGRVHITGATLALLAGAYAVEDAGMEH 431

Qy        630 ----------PGF-VFTPRSREE 641
                       | : |  ||: ||
Db        432 RDPYLRELGEPTYLVIDPRAEEE 454


RESULT 3
US-08-726-214-8
; Sequence 8, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1064 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-8

  Query Match             9.0%;  Score 322;  DB 3;  Length 1064;
  Best Local Similarity   31.2%;  Pred. No. 1e-22;
  Matches   84;  Conservative   50;  Mismatches   89;  Indels   46;  Gaps    7;

Qy        397 GLYLSDIPIHNALRDVV-------------LIGEQARAQDGLKKRLGKLKATLEQA--- 439
              ||||   |   : ||::              ::|   :|  | :|   |:||  :|
Db        157 GLYLGWRP--ESQRDLLPQLAANAVLFLCGNVVGAYHKA---LMER--ALRATFREALSS 209

Qy        440 ---HQALEEEKKKTVDLLCSIFP--------CEVAQQLWQGQ-----------VVQAKKF 477
                  : |: |||    || || |          |:  :|  ||           :  |:
Db        210 LHSRRRLDTEKKHQEHLLLSILPAYLAREMKAEIMARLQAGQSSRPENTNNFHSLYVKRH 269

Qy        478 SNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGG 537
               |::|::|||||| : |:||| ::: ||| |: :|||   | :  ::: :|| |    |
Db        270 QGVSVLYADIVGFTRLASECSPKELVLMLNELFGKFDQIAKEHECMRIKILGDCYYCVSG 329

Qy        538 LHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLF 597
              |      ||:    | | |     ::    |  | ||:|:|||||  ||:|::  :| ::
Db        330 LPLSLPDHAINCVRMGLDMCRAIRKLRVATGVDINMRVGVHSGSVLCGVIGLQKWQYDVW 389

Qy        598 GNNVTLANKFESCSVPRKINVSPTTYRLL 626
               ::|||'||  |:  || :::::  |  ||
Db        390 SHDVTLANHMEAGGVPGRVHITGATLALL 418


RESULT 4
US-08-726-214-14
; Sequence 14, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
```

```
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1099 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-14

  Query Match             8.9%;  Score 319.5;  DB 3;  Length 1099;
  Best Local Similarity   26.3%;  Pred. No. 1.9e-22;
  Matches   82;  Conservative   64;  Mismatches   95;  Indels   71;  Gaps    8;

Qy        376 SSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKAT 435
              ::|:: ||        :|||         : :| ||: :                    |
Db        184 ANAVILLGG-------NFTGA---FHKHQLQDASRDLFIY-------------------T 214

Qy        436 LE--QAHQALEEEKKKTVDLLCSIFPCEVA--------QQLWQG-----------QVVQA 474
              ::  |  : |  ||::  :|| |: |  ::          ::| :|           :
Db        215 VKCIQIRRKLRVEKRQQENLLLSVLPAHISMGMKLAIIERLKEGGDRHYMPDNNFHSLYV 274

Qy        475 KKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCV 534
              |:  ||::|::|||||| : | ||| ::: :|| |: :|||     :  ::: :|| |
Db        275 KRHQNVSILYADIVGFTRLASDCSPKELVVVLNELFGKFDQIAKANECMRIKILGDCYYC 334

Qy        535 AGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRY 594
              ||     |||     | | : |   :|    |  | ||:|:|||:|  ||:|::  :|
Db        335 VSGLPVSLPTHARNCVKMGLDICEAIKQVREATGVDISMRVGIHSGNVLCGVIGLRKWQY 394

Qy        595 CLFGNNVTLANKFESCSVPRKINVSPTTYRLL------KDCPG--------------FV 633
              :: ::|:|||: |:  || :::::  |   |       :|  |              |
Db        395 DVWSHDVSLANRMEAAGVPGRVHITEATLNHLDKAYEVEDGHGEQRDPYLKEMNIRTYLV 454

Qy        634 FTPRSREELPPN 645
              |||::  ||:
Db        455 IDPRSQQPPPPS 466


RESULT 5
US-08-307-896-3
; Sequence 3, Application US/08307896C
; Patent No. 6034071
; GENERAL INFORMATION:
;  APPLICANT: Iyengar, Srinivas Ravi
;  TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND ADENYLYL
;  TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
;  FILE REFERENCE: 29770
;  CURRENT APPLICATION NUMBER: US/08/307,896C
;  CURRENT FILING DATE: 1994-09-16
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1090
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-08-307-896-3

  Query Match             8.9%;  Score 319;  DB 3;  Length 1090;
  Best Local Similarity   30.4%;  Pred. No. 2.1e-22;
  Matches   80;  Conservative   53;  Mismatches   86;  Indels   44;  Gaps    7;

Qy        398 LYLSDIP------IHNALRDVVLI--GEQARAQDGLKKRLGKLKATLEQAHQ-------- 441
              :|||  |          | :|::   |  | |     | |  ::  |:| ::
Db        174 VYLSATPGAKEHLFWQILANVIIFICGNLAGA---YHKHL--MELALQQTYRDTCNCIKS 228

Qy        442 --ALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQ-------------------AKKFSN 479
                || ||::   || |: |  :| :: : :::|                   |: :|
Db        229 RIKLEFEKRQQERLLLSLLPAHIAMEM-KAEIIQRLQGPKAGQMENTNNFHNLYVKRHTN 287

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539
              |::|::|||||| : | ||| ::: ||| |: :|||   | :  ::: :|| |    ||
Db        288 VSILYADIVGFTRLASDCSPGELVHMLNELFGKFDQIAKENECMRIKILGDCYYCVSGLP 347

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
                ||     | | | |   :|    |  | ||:|:|||:|  ||:|::  :| :: :
Db        348 ISLPNHAKNCVKMGLDMCEAIKKVRDATGVDINMRVGVHSGNVLCGVIGLQKWQYDVWSH 407

Qy        600 NVTLANKFESCSVPRKINVSPTT 622
              :|||||  |:  || ::::|  |
Db        408 DVTLANHMEAGGVPGRVHISSVT 430


RESULT 6
US-08-726-214-4
; Sequence 4, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1090 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-4

  Query Match             8.9%;  Score 319;  DB 3;  Length 1090;
```

```
  Best Local Similarity   30.4%;  Pred. No. 2.1e-22;
  Matches   80;  Conservative   53;  Mismatches   86;  Indels   44;  Gaps    7;

Qy        398 LYLSDIP------IHNALRDVVLI--GEQARAQDGLKKRLGKLKATLEQAHQ-------- 441
              :|||  |          | :|::   |  | |     | |  ::  |:| ::
Db        174 VYLSATPGAKEHLFWQILANVIIFICGNLAGA---YHKHL--MELALQQTYRDTCNCIKS 228

Qy        442 --ALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQ-------------------AKKFSN 479
                 || ||::   || |: |  :| :: : :::|                    |: :|
Db        229 RIKLEFEKRQQERLLLSLLPAHIAMEM-KAEIIQRLQGPKAGQMENTNNFHNLYVKRHTN 287

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539
              |::|::|||||| : | ||| ::: ||| |: :|||   | :  ::: :|| |    ||
Db        288 VSILYADIVGFTRLASDCSPGELVHMLNELFGKFDQIAKENECMRIKILGDCYYCVSGLP 347

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
                 ||     | | | |   :|    |  | ||:|:|||:|  ||:|::  :| :: :
Db        348 ISLPNHAKNCVKMGLDMCEAIKKVRDATGVDINMRVGVHSGNVLCGVIGLQKWQYDVWSH 407

Qy        600 NVTLANKFESCSVPRKINVSPTT 622
              :|||||  |:  || ::::|  |
Db        408 DVTLANHMEAGGVPGRVHISSVT 430


RESULT 7
PCT-US95-11808-3
; Sequence 3, Application PC/TUS9511808
;  GENERAL INFORMATION:
;    APPLICANT:  Iyengar, Srinivas Ravi V.
;    TITLE OF INVENTION:  MUTANT ACTIVATED GsALPHA AND
;    TITLE OF INVENTION:  ADENYLYL
;    TITLE OF INVENTION:  CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brumbaugh, Graves, Donohue and
;      ADDRESSEE:  Raymond
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.
;      ZIP:  10112-0228
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/11808
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,896
;      FILING DATE:  16-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Richard S.
;      REGISTRATION NUMBER:  26,154
;      REFERENCE/DOCKET NUMBER:  29970 165/28755
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 408-2500
;      TELEFAX:  (212) 765-2519
;      TELEX:  650 6111063
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1090 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-11808-3

  Query Match             8.9%;  Score 319;  DB 5;  Length 1090;
  Best Local Similarity   30.4%;  Pred. No. 2.1e-22;
  Matches   80;  Conservative   53;  Mismatches   86;  Indels   44;  Gaps    7;

Qy        398 LYLSDIP------IHNALRDVVLI--GEQARAQDGLKKRLGKLKATLEQAHQ-------- 441
              :|||  |          | :|::   |  | |     | |  ::  |:| ::
Db        174 VYLSATPGAKEHLFWQILANVIIFICGNLAGA---YHKHL--MELALQQTYRDTCNCIKS 228

Qy        442 --ALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQ-------------------AKKFSN 479
                 || ||::   || |: |  :| :: : :::|                    |: :|
Db        229 RIKLEFEKRQQERLLLSLLPAHIAMEM-KAEIIQRLQGPKAGQMENTNNFHNLYVKRHTN 287

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539
              |::|::|||||| : | ||| ::: ||| |: :|||   | :  ::: :|| |    ||
Db        288 VSILYADIVGFTRLASDCSPGELVHMLNELFGKFDQIAKENECMRIKILGDCYYCVSGLP 347

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
                 ||     | | | |   :|    |  | ||:|:|||:|  ||:|::  :| :: :
Db        348 ISLPNHAKNCVKMGLDMCEAIKKVRDATGVDINMRVGVHSGNVLCGVIGLQKWQYDVWSH 407

Qy        600 NVTLANKFESCSVPRKINVSPTT 622
              :|||||  |:  || ::::|  |
Db        408 DVTLANHMEAGGVPGRVHISSVT 430


RESULT 8
US-08-726-214-6
; Sequence 6, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1144 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-6

  Query Match             8.4%;  Score 303;  DB 3;  Length 1144;
  Best Local Similarity   30.7%;  Pred. No. 8.8e-21;
  Matches   84;  Conservative   50;  Mismatches   92;  Indels   48;  Gaps    9;

Qy        433 KATLE-----QAHQALEEEKKKTVDLLCSIFPCEVAQQL------------WQGQVVQA 474
              || ||      :    |||: ::  :|: || |  || ::            |   :
Db        253 KAFLEARQSLEVKMNLEEQSQQQENLMLSILPKHVADEMLKDMKKDESQKDQQQFNTMYM 312
```

```
Qy        475 KKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCV 534
              :  ||::||:|||||| : | ||  ::: :|| |: |||:  :    ::: :|| |
Db        313 YRHENVSILFADIVGFTQLSSACSAQELVKLLNELFARFDKLAAKYHQLRIKILGDCYYC 372

Qy        535 AGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRY 594
               ||    : |||   || | |:|    |       : ||:|:|:|:|  ||:| |  :|
Db        373 ICGLPDYREDHAVCSILMGLAMVEAISYVREKTKTGVDMRVGVHTGTVLGGVLGQKRWQY 432

Qy        595 CLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG-- 652
              ::  :||:||| |:  :| ::::| :|     ||           |   |  | ||
Db        433 DVWSTDVTVANKMEAGGIPGRVHISQST----MDC-----------LKGEFDVE-PGDG 475

Qy        653 --ICHFLDAYQQGTN------SKPCFQKKDVEDG 678
                | :||  ::|       |||   ||  ::|
Db        476 GSRCDYLD--EKGIETYLIIASKP-EVKKTAQNG 506


RESULT 9
US-08-726-214-2
; Sequence 2, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1134 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-2

  Query Match             8.4%;  Score 302;  DB 3;  Length 1134;
  Best Local Similarity   28.1%;  Pred. No. 1.1e-20;
  Matches   73;  Conservative   57;  Mismatches   94;  Indels   36;  Gaps    4;

Qy        376 SSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKAT 435
              ::|:||||                  : ::    | :: |:|: :  |: |
Db        218 ANALLFLG------------------VNVYGIF--VRILAERAQRKAFLQAR------N 250

Qy        436 LEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG---------QVVQAKKFSNVTMLFSD 486
                :   ||:| :|   || |: |  || :: :           :  ::  ||::||:|
Db        251 CIEDRLRLEDENEKQERLLMSLLPRNVAMEMKEDFLKPPERIFHKIYIQRHDNVSILFAD 310

Qy        487 IVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHA 546
              |||||  : |||:  ::: :|| |: :|||:    |    ::: :|| |    || :    ||
Db        311 IVGFTGLASQCTAQELVKLLNELFGKFDELATENHCRRIKILGDCYYCVSGLTQPKTDHA 370

Qy        547 VQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANK 606
                | | |::    |       : ||:|||:| |  ||:|::  :| :: |:|||||
Db        371 HCCVEMGLDMIDTITSVAEATEVDLNMRVGLHTGRVLCGVLGLRKWQYDVWSNDVTLANV 430

Qy        607 FESCSVPRKINVSPTTYRLL 626
              |:  :| |:::: ||   |
Db        431 MEAAGLPGKVHITKTTLACL 450


RESULT 10
US-08-726-214-18
; Sequence 18, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  533 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-18

  Query Match             8.3%;  Score 299.5;  DB 3;  Length 533;
  Best Local Similarity   32.6%;  Pred. No. 5.5e-21;
  Matches   63;  Conservative   44;  Mismatches   77;  Indels    9;  Gaps    1;

Qy        443 LEEEKKKTVDLLCSIFPCEVAQQLWQG---------QVVQAKKFSNVTMLFSDIVGFTAI 493
              ||:| :|   || |: |  || :: :           :  ::  ||::||:|||||| :
Db         23 LEDENEKQERLLMSLLPRNVAMEMKEDFLKPPERIFHKIYIQRHDNVSILFADIVGFTGL 82

Qy        494 CSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALMA 553
              |||:  ::: :|| |: :|||:    |    ::: :|| |    || :    ||     |
Db         83 ASQCTAQELVKLLNELFGKFDELATENHCRRIKILGDCYYCVSGLTQPKTDHAHCCVEMG 142

Qy        554 LKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVP 613
              | |::    |       : ||:|||:| |  ||:|::  :| :: |:|||||  |:  :|
Db        143 LDMIDTITSVAEATEVDLNMRVGLHTGRVLCGVLGLRKWQYDVWSNDVTLANVMEAAGLP 202
```

```
Qy        614 RKINVSPTTYRLL 626
              |:::: ||   |
Db        203 GKVHITKTTLACL 215


RESULT 11
US-08-864-785-2
; Sequence 2, Application US/08864785A
; Patent No. 6329566
; GENERAL INFORMATION:
;  APPLICANT: Kaplan, Joshua M.
;  APPLICANT: Oppenheimer, Allison J.
;  APPLICANT: Hart, Anne C.
;  TITLE OF INVENTION: METHODS FOR THE DETECTION, TREATMENT,
;  TITLE OF INVENTION: AND PREVENTION OF NEURODEGENERATION
;  FILE REFERENCE: 00786/353001
;  CURRENT APPLICATION NUMBER: US/08/864,785A
;  CURRENT FILING DATE: 1997-05-29
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1253
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-08-864-785-2

  Query Match             8.0%;  Score 287.5;  DB 4;  Length 1253;
  Best Local Similarity   30.5%;  Pred. No. 3.6e-19;
  Matches   69;  Conservative   42;  Mismatches   78;  Indels   37;  Gaps    2;

Qy        439 AHQALEEEKKKTVDLLCSIFPCEVAQQLWQG---------------------------- 469
              | : || | :    :: |: | :|| :| :
Db        240 ARKDLELETQFKDHMIQSVMPKKVADELLKDASELRRPSASNDSNCRTSNATQVDQPLAK 299

Qy        470 QVVQAKKF--------SNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELD 521
              | : :||        :||::||:|| ||| : |  |  ::: :|| |: |||  |
Db        300 MVPEYRKFRPFTMNLMTNVSILFADIAGFTKMSSNKSADELVNLLNDLFGRFDTLCRLRG 359

Qy        522 VYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGS 581
              : |: |:|| |    |  :  | || :   | | |:    :     |: : ||:|:|:|
Db        360 LEKISTLGDCYYCVAGCPEPCDDHACRTVEMGLDMIVAIRQFDIDRGQEVNMRVGIHTGK 419

Qy        582 VFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLK 627
              |  |:|| |  :: :| |:|||||: ||  |  :::||  | :|||
Db        420 VMCGMVGTKRFKFDVFSNDVTLANEMESSGVAGRVHVSEATAKLLK 465


RESULT 12
US-08-726-214-16
; Sequence 16, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1248 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-16

  Query Match             7.9%;  Score 284.5;  DB 3;  Length 1248;
  Best Local Similarity   22.5%;  Pred. No. 7e-19;
  Matches  147;  Conservative   99;  Mismatches  217;  Indels  189;  Gaps   31;

Qy        103 KESRKSLEREDFEKTIAEQAVAA--GVPVEVIKESLGEEVFKICYEED---------- 148
              ||  : | : : |  : :|   :   :| :::|||:       | :
Db        568 KERNEFLRKHNIETYLIKQPEESLLSLPEDIVKESVS------CSDRRNSGATFTEGSWS 621

Qy        149 -----ENILG---VVGGTLKDFLNSFSTLLKQSSHCQ---------------EAGKRGR 184
                  :||:|    :    :: :|    | |: | |                 :| : |
Db        622 PELPFDNIVGKQNTLAALTRNSINLLPNHLAQALHVQSGPEEINKRIEHTIDLRSGDKLR 681

Qy        185 LEDASILCLDKEDDFL-HVY----------------YFFPKRTTSLILPG-------- 217
              |     |  :|  | | |                 |    || ||
Db        682 REHIKPFSLMFKDSSLEHKYSQMRDEVFKSNLVCAFIVLLFITAIQSL-LPSSRLMPMTI 740

Qy        218 -----IIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQ 272
                  |:  :| ||  |  :   :|      |  ::|: ||  :| : :
Db        741 QFSILIMLHSALVLITTAEDYKCLPLILRKTCC-WINETYLARNVIIFA---------- 788

Qy        273 SSLV-----IPTSLFCKTFPFHFMFDKDMTI--LQFGNGIRRLMNRRDFQGKPNFEEYFE 325
              | |:     :   |:|       ||| : :  | | :          |    :: |||
Db        789 SILINFLGAVINILWCD-------FDKSIPLKNLTFNSS-------AVFTDICSYPEYF- 833

Qy        326 ILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAI---LFL 382
                    |:|:: |:     :|:    ||| | : :: :   || : :  |   |||
Db        834 --------VFTGVLAMVTCAVFLRL----NSVLKLAVLLIMIA--IYALLTETIYAGLFL 879

Qy        383 GSPCVDRL----EDFTG--------RGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430
              | |    ||| |           ::|  :  |          |:|
Db        880 S---YDNLNHSGEDFLGTKEASLLLMAMFLLAVFYH---------GQQL---------- 916

Qy        431 KLKATLE-----QAHQALEEEK---KKTVDLLCSIFPCEVAQQLWQ----GQVVQAKKFS 478
              :  | |:     || : : | |   :   ::| :| |  ||:  :     : : :: :
Db        917 EYTARLDFLWRVQAKEEINEMKDLREHNENMLRNILPGHVARHFLEKDRDNEELYSQSYD 976

Qy        479 NVTMLFSDIVGFTAICSQC----SPLQVITMLNALYTRFDQQCGE---LDVYKVETIGDA 531
               | ::|: | ||    ||       :: : :|| :   ||:  ||    |: |::|||
Db        977 AVGVMFASIPGFADFYSQTEMNNQGVECLRLLNEIIADFDELLGEDRFQDIEKIKTIGST 1036

Qy        532 YCVAGGLHKESDT------HAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAG 585
              |    ||  |        |   :|  :| : |   |:        ::|||:  ||| ||
Db       1037 YMAVSGLSPEKQQCEDKWGHLCALADFSLALTESIQEINKHSFNNFELRIGISHGSVVAG 1096

Qy        586 VVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPR 637
              |:| | |:| ::|  | ||:: :|  |  :| |   || :|||  || |  |
Db       1097 VIGAKKPQYDIWGKTVNLASRMDSTGVSGRIQVPEETYLILKD-QGFAFDYR 1147


RESULT 13
US-08-726-214-10
; Sequence 10, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
```

```
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1098 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-10

  Query Match             7.9%;  Score 283.5;  DB 3;  Length 1098;
  Best Local Similarity   26.4%;  Pred. No. 7.1e-19;
  Matches   83;  Conservative   52;  Mismatches  108;  Indels   71;  Gaps    8;

Qy        438 QAHQALEEEKKKTVDLLCSIFPCEVAQQL----------WQGQVVQAKKFSNVTMLFSDI 487
              ||    : | ::   || |: |  || ::          :  :|  ||::||:||
Db        253 QARLHSQRENQQQERLLLSVLPRHVAMEMKADINAKQEDMMFHKIYIQKHDNVSILFADI 312

Qy        488 VGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAV 547
               |||:: |||:  :::  || |: |||:   |    ::: :|| |    || :    ||
Db        313 EGFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEARADHAH 372

Qy        548 QIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKF 607
               | : |:|    |    |  : ||:|:||| |  ||:|::  :: :: |:|||||
Db        373 CCVEMGMDMIEAISSVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHM 432

Qy        608 ESCSVPRKINVSPTT-------YRLLKDCPG-------------FVF---TPRSREEL-- 642
              |:     :|:::  |       | :  | |             |:    | : :||
Db        433 EAGGKAGRIHITKATLNYLNGDYEVEPGCGGERNAYLKEHSIETFLILRCTQKRKEEKAM 492

Qy        643 --------------PPNFPSEIPGICHFLDAYQQGTNSKPCFQKK----DVEDGNA--- 680
                           ||:: :| |   |       | |       |:    | :| ||
Db        493 IAKMNRQRTNSIGHNPPHWGAERPFYNHL------GGNQVSKEMKRMGFEDPKDKNAQES 546

Qy        681 --------NFLGKA 686
                      |||:|
Db        547 ANPEDEVDEFLGRA 560


RESULT 14
US-09-008-097-4
; Sequence 4, Application US/09008097
; Patent No. 6306830
;  GENERAL INFORMATION:
;    APPLICANT:  Hammond, H. Kirk
;    APPLICANT:  Insel, Paul A.
;    APPLICANT:  Ping,.Peipei
;    APPLICANT:  Post, Steven R.
;    APPLICANT:  Gao, Meihua
;    TITLE OF INVENTION:  GENE THERAPY FOR CONGESTIVE
;    TITLE OF INVENTION:  HEART FAILURE
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/008,097
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dylan, Tyler M
;      REGISTRATION NUMBER:  37,612
;      REFERENCE/DOCKET NUMBER:  22000-20567.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  604 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-008-097-4

  Query Match             7.9%;  Score 282.5;  DB 4;  Length 604;
  Best Local Similarity   31.8%;  Pred. No. 3.3e-19;
  Matches   63;  Conservative   44;  Mismatches   82;  Indels    9;  Gaps    1;

Qy        438 QAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG---------QVVQAKKFSNVTMLFSDIV 488
              ||    |: | ::   || |: |  || :: :           :  :|  ||::||:||
Db        115 QARLHLQHENRQQERLLLSVLPQHVAMEMKEDINTKKEDMFHKIYIQKHDNVSILFADIE 174

Qy        489 GFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQ 548
              |||:: |||:  :::  || |: |||:   |    ::: :|| |    || :    ||
Db        175 GFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEARADHAHC 234

Qy        549 IALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFE 608
               | : |:|    |    |  : ||:|:||| |  ||:|::  :: :: |:|||||  |
Db        235 CVEMGVDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHME 294

Qy        609 SCSVPRKINVSPTTYRLL 626
              : |   :|:::  | : |
Db        295 AGSRAGRIHITRATLQYL 312


RESULT 15
US-09-008-097-6
; Sequence 6, Application US/09008097
; Patent No. 6306830
;  GENERAL INFORMATION:
;    APPLICANT:  Hammond, H. Kirk
;    APPLICANT:  Insel, Paul A.
;    APPLICANT:  Ping, Peipei
```

```
;    APPLICANT:  Post, Steven R.
;    APPLICANT:  Gao, Meihua
;    TITLE OF INVENTION:  GENE THERAPY FOR CONGESTIVE
;    TITLE OF INVENTION:  HEART FAILURE
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/008,097
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dylan, Tyler M
;      REGISTRATION NUMBER:  37,612
;      REFERENCE/DOCKET NUMBER:  22000-20567.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1167 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-008-097-6

  Query Match             7.9%;  Score 282.5;  DB 4;  Length 1167;
  Best Local Similarity   31.8%;  Pred. No. 9.9e-19;
  Matches   63;  Conservative   44;  Mismatches   82;  Indels    9;  Gaps    1;

Qy        438 QAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG---------QVVQAKKFSNVTMLFSDIV 488
              ||   |: | ::   || |: |  || :: :             :  :|  ||::||:||
Db        326 QARLHLQHENRQQERLLLSVLPQHVAMEMKEDINTKKEDMFHKIYIQKHDNVSILFADIE 385

Qy        489 GFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQ 548
              |||:: |||:  :::  || |: |||:   |    ::: :|| |    || :    ||
Db        386 GFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEARADHAHC 445

Qy        549 IALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFE 608
                | : |:|    |    |  : ||:|:||| |  ||:|::  :: :: |:|||||  |
Db        446 CVEMGVDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHME 505

Qy        609 SCSVPRKINVSPTTYRLL 626
              : |   :|:::  | : |
Db        506 AGSRAGRIHITRATLQYL 523


Search completed: June 27, 2003, 13:05:57
Job time : 16.2322 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:05:32 ; Search time 36.3713 Seconds
(without alignments)
2080.112 Million cell updates/sec

Title: US-09-762-767A-2
Perfect score: 3593
Sequence: 1 MFCTKLKDLKITGECPFSLL.....QKKDVEDGNANFLGKASGID 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3593	100.0	690	9	US-10-205-823-162	Sequence 162, App
2	3265.5	90.9	691	9	US-09-952-213D-2	Sequence 2, Appli
3	408.5	11.4	1073	9	US-10-157-031-18	Sequence 18, Appl
4	408.5	11.4	1073	10	US-09-819-249-2	Sequence 2, Appli
5	325.5	9.1	1077	9	US-10-121-911-1	Sequence 1, Appli
6	316	8.8	1080	9	US-10-282-942-2	Sequence 2, Appli
7	303.5	8.4	855	10	US-09-925-297-811	Sequence 811, App
8	282.5	7.9	604	10	US-09-750-240-4	Sequence 4, Appli
9	282.5	7.9	1167	10	US-09-750-240-6	Sequence 6, Appli
10	280.5	7.8	1261	9	US-10-175-158-2	Sequence 2, Appli
11	279.5	7.8	1168	9	US-10-201-000-2	Sequence 2, Appli
12	279.5	7.8	1168	10	US-09-750-240-11	Sequence 11, Appl
13	278	7.7	1167	10	US-09-750-240-13	Sequence 13, Appl
14	263.5	7.3	1294	12	US-10-071-223-2	Sequence 2, Appli
15	260.5	7.3	1353	10	US-09-751-100B-2	Sequence 2, Appli
16	260.5	7.3	1353	12	US-10-071-223-3	Sequence 3, Appli
17	255.5	7.1	1353	10	US-09-751-100B-99	Sequence 99, Appl
18	212	5.9	257	9	US-09-989-442-95	Sequence 95, Appl
19	207	5.8	241	9	US-09-764-868-794	Sequence 794, App
20	207	5.8	241	9	US-09-989-442-141	Sequence 141, App
21	195	5.4	670	9	US-10-051-643-178	Sequence 178, App
22	195	5.4	670	9	US-09-880-505-178	Sequence 178, App
23	195	5.4	722	9	US-10-051-643-174	Sequence 174, App
24	195	5.4	722	9	US-09-880-505-174	Sequence 174, App
25	194	5.4	302	10	US-09-915-582-50	Sequence 50, Appl
26	192.5	5.4	107	10	US-09-205-658-76	Sequence 76, Appl
27	192.5	5.4	107	10	US-09-844-353A-76	Sequence 76, Appl
28	191	5.3	112	10	US-09-205-658-75	Sequence 75, Appl
29	191	5.3	112	10	US-09-844-353A-75	Sequence 75, Appl
30	187	5.2	419	9	US-10-051-643-125	Sequence 125, App
31	187	5.2	419	9	US-09-880-505-125	Sequence 125, App
32	181.5	5.1	311	10	US-09-925-297-515	Sequence 515, App
33	174	4.8	225	9	US-09-989-442-92	Sequence 92, Appl
34	148	4.1	109	9	US-09-764-868-795	Sequence 795, App
35	148	4.1	109	9	US-09-989-442-143	Sequence 143, App
36	148	4.1	109	9	US-10-091-504-1060	Sequence 1060, Ap
37	148	4.1	109	10	US-09-764-869-1060	Sequence 1060, Ap
38	138.5	3.9	85	10	US-09-751-100B-59	Sequence 59, Appl
39	137.5	3.8	85	10	US-09-751-100B-60	Sequence 60, Appl
40	137.5	3.8	85	10	US-09-751-100B-61	Sequence 61, Appl
41	133.5	3.7	85	10	US-09-751-100B-54	Sequence 54, Appl
42	133.5	3.7	85	10	US-09-751-100B-55	Sequence 55, Appl
43	133.5	3.7	85	10	US-09-751-100B-56	Sequence 56, Appl
44	133.5	3.7	85	10	US-09-751-100B-57	Sequence 57, Appl
45	133.5	3.7	85	10	US-09-751-100B-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-10-205-823-162
; Sequence 162, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-162

Query Match 100.0%; Score 3593; DB 9; Length 690;
Best Local Similarity 100.0%; Pred. No. 1e-305;

Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660

Qy        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
              |||||||||||||||||||||||||||||
Db        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
```

RESULT 2
US-09-952-213D-2
; Sequence 2, Application US/09952213D
; Publication No. US20030096240A1
; GENERAL INFORMATION:
; APPLICANT: MURAD, FERID
; APPLICANT: SHARINA, IRAIDA G.
; APPLICANT: KRUMENACKER, J. S.
; APPLICANT: MARTIN, E.
; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC
; FILE REFERENCE: UTSH:252US
; CURRENT APPLICATION NUMBER: US/09/952,213D
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-952-213D-2

Query Match 90.9%; Score 3265.5; DB 9; Length 691;

Best Local Similarity 89.6%; Pred. No. 4.8e-277;
Matches 619; Conservative 37; Mismatches 34; Indels 1; Gaps 1;

```
Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||| | |||||||||||||||||||| | :|| || || |:||:|||| |||| : |||
Db          1 MFCRKFKDLKITGECPFSLLAPGQVPKEPTEEVAGGSEGCQATLPICQYFPEKNAEGSLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||:|||||||||||||||||| ||||:||||||||||:|:||| |:|| || |||
Db         61 QRKTSRNRVYLHTLAESICKLIFPECERLNLALQRTLAKHKIEENRKSSEKEDLEKIIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              :|:||| ||| :|:|||||:|||||||||:||||||||||||||||||||||||||||||
Db        121 EAIAAGAPVEALKDSLGEELFKICYEEDEHILGVVGGTLKDFLNSFSTLLKQSSHCQEAE 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              :|||||||||||||: |||:||||||||||:|:|||||||| :|||: |||||||||
Db        181 RRGRLEDASILCLDKDQDFLNVYYFFPKRTTALLLPGIIKAAARILYESHVEVSLMPPCF 240

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              :||:|||||||||||||:|||||||| |||||||| |||||||||||| |:|: |||
Db        241 RSDCTEFVNQPYLLYSVHVKSTKPSLSPGKPQSSLVIPASLFCKTFPFHFMLDRDLAILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              |||||||:|:|||||||||||:||||||||||||||||||||:|||||||||||||||||
Db        301 LGNGIRRLVNKRDFQGKPNFEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              |||||||||||||||| |||||||||:|||||||||| |||||||||:||||||| |
Db        421 AQDGLKKRLGKLKATLEHAHQALEEEKKRTVDLLCSIFPSEVAQQLWQGQIVQAKKFSEV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHR 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLDAY 660

Qy        661 -QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
               || |||| || |||||||||||||||:|
Db        661 HHQGPNSKPWFQDKDVEDGNANFLGKASGVD 691
```

RESULT 3
US-10-157-031-18
; Sequence 18, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequenc
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-18

```
  Query Match             11.4%;  Score 408.5;  DB 9;  Length 1073;
  Best Local Similarity   37.3%;  Pred. No. 1.2e-26;
  Matches  101;  Conservative   43;  Mismatches   92;  Indels   35;  Gaps    6;

Qy        423 DGLKKRL----GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFS 478
              | | :||      |:  :|:  |  : |: :   |   : |  | : | :   |: : :
Db        762 DTLIRRLQLYSRNLEHLVEERTQLYKAERDRADRLNFMLLPRLVVKSLKEKGFVEPELYE 821

Qy        479 NVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL 538
              ||: |||||||| ||   :|:::|: ||| :|  ||      ||||||||||| || ||
Db        822 EVTIYFSDIVGFTTICKYSTPMEVVDMLNDIYKSFDHIVDHHDVYKVETIGDAYMVASGL 881

Qy        539 HKES-DTHAVQIALMALKMMELSDEVMSPH--GEPIKMRIGLHSGSVFAGVVGVKMPRYC 595
              | : : ||: || |||:::         |  | || :|||:|||   |||||:||||||
Db        882 PKRNGNRHAIDIAKMALEILSFMGTFELEHLPGLPIWIRIGVHSGPCAAGVVGIKMPRYC 941

Qy        596 LFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICH 655
              |||: |  |:: ||  :| :|:|| :|  :||          | |           |
Db        942 LFGDTVNTASRMESTGLPLRIHVSGSTIAILK----------RTE----------CQ 978

Qy        656 FL-----DAYQQGTNSKPCFQKKDVEDGNAN 681
              ||     : | :|  ::  :    ::|   |
Db        979 FLYEVRGETYLKGRGNETTYWLTGMKDQKFN 1009


RESULT 4
US-09-819-249-2
; Sequence 2, Application US/09819249
; Patent No. US20010029019A1
; GENERAL INFORMATION:
;  APPLICANT: Waldman, Scott A.
;  APPLICANT:  Park, Jason
;  APPLICANT:  Schulz, Stephanie
;  TITLE OF INVENTION: Compositions And Methods For Identifying And Targeting Cancer Cel
;  TITLE OF INVENTION:  Alimentary Canal Origin
;  FILE REFERENCE: TJU2412
;  CURRENT APPLICATION NUMBER: US/09/819,249
;  CURRENT FILING DATE:  2001-03-27
;  PRIOR APPLICATION NUMBER: 60/192,229
;  PRIOR FILING DATE: 2000-03-27
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 1073
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-819-249-2

  Query Match             11.4%;  Score 408.5;  DB 10;  Length 1073;
  Best Local Similarity   37.3%;  Pred. No. 1.2e-26;
  Matches  101;  Conservative   43;  Mismatches   92;  Indels   35;  Gaps    6;

Qy        423 DGLKKRL----GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFS 478
              | | :||      |:  :|:  |  : |: :   |   : |  | : | :   |: : :
Db        762 DTLIRRLQLYSRNLEHLVEERTQLYKAERDRADRLNFMLLPRLVVKSLKEKGFVEPELYE 821

Qy        479 NVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL 538
              ||: |||||||| ||   :|:::|: ||| :|  ||      ||||||||||| || ||
Db        822 EVTIYFSDIVGFTTICKYSTPMEVVDMLNDIYKSFDHIVDHHDVYKVETIGDAYMVASGL 881

Qy        539 HKES-DTHAVQIALMALKMMELSDEVMSPH--GEPIKMRIGLHSGSVFAGVVGVKMPRYC 595
              | : : ||: || |||:::         |  | || :|||:|||   |||||:||||||
Db        882 PKRNGNRHAIDIAKMALEILSFMGTFELEHLPGLPIWIRIGVHSGPCAAGVVGIKMPRYC 941

Qy        596 LFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICH 655
              |||: |  |:: ||  :| :|:|| :|  :||          | |           |
Db        942 LFGDTVNTASRMESTGLPLRIHVSGSTIAILK----------RTE----------CQ 978

Qy        656 FL-----DAYQQGTNSKPCFQKKDVEDGNAN 681
              ||     : | :|  ::  :    ::|   |
Db        979 FLYEVRGETYLKGRGNETTYWLTGMKDQKFN 1009


RESULT 5
US-10-121-911-1
; Sequence 1, Application US/10121911
; Patent No. US20020164632A1
; GENERAL INFORMATION:
;  APPLICANT: Kapeller-Libermann, Rosana
;  TITLE OF INVENTION: 21529, A NOVEL ADENYLATE CYCLASE
;  FILE REFERENCE: 5800-47
;  CURRENT APPLICATION NUMBER: US/10/121,911
;  CURRENT FILING DATE:  2002-04-12
;  PRIOR APPLICATION NUMBER: US/09/412,210
;  PRIOR FILING DATE: 1999-10-05
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1077
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-121-911-1

  Query Match             9.1%;  Score 325.5;  DB 9;  Length 1077;
  Best Local Similarity   28.2%;  Pred. No. 2.2e-19;
  Matches   91;  Conservative   55;  Mismatches   98;  Indels   79;  Gaps    9;

Qy        365 DLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDG 424
              | :  ::  : ::|:|||               |:|      | || :        ||
Db        165 DSRPALLPQLAANAVLFLCGNVA---------GVY------HKALME--------RA--- 198

Qy        425 LKKRLGKLKATLEQA------HQALEEEKKKTVDLLCSIFP--------CEVAQQLWQGQ 470
                    |:||  :|       : |: |||    || || |         |:  :|  ||
Db        199 -------LRATFREALSSLHSRRRLDTEKKHQEHLLLSILPAYLAREMKAEIMARLQAGQ 251

Qy        471 ----------VVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGE 519
                       :  |:   |::|::|||||| : |:||| ::: ||| |: :|||   |
Db        252 GSRPESTNNFHSLYVKRHQGVSVLYADIVGFTRLASECSPKELVLMLNELFGKFDQIAKE 311

Qy        520 LDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHS 579
              :  ::: :|| |    ||      ||:    | | |     :: :  |  | ||:|:||
Db        312 HECMRIKILGDCYYCVSGLPLSLPDHAINCVRMGLDMCRAIRKLRAATGVDINMRVGVHS 371

Qy        580 GSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDC---------- 629
              |||  ||:|::  :| :: ::|||||  |:  || :::::  |  ||
Db        372 GSVLCGVIGLQKWQYDVWSHDVTLANHMEAGGVPGRVHITGATLALLAGAYAVEDAGMEH 431

Qy        630 ----------PGF-VFTPRSREE 641
                       | : |  ||: ||
Db        432 RDPYLRELGEPTYLVIDPRAEEE 454


RESULT 6
US-10-282-942-2
; Sequence 2, Application US/10282942
; Publication No. US20030087295A1
; GENERAL INFORMATION:
;  APPLICANT: Silos-Santiago, Inmaculada
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
;  TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF PAIN DISORDERS USING 9805
;  FILE REFERENCE: MPI01-240P1RM
;  CURRENT APPLICATION NUMBER: US/10/282,942
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: 60/335,047
;  PRIOR FILING DATE: 2001-10-31
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1080
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-282-942-2

  Query Match             8.8%;  Score 316;  DB 9;  Length 1080;
  Best Local Similarity   28.2%;  Pred. No. 1.5e-18;
```

```
  Matches   70;  Conservative   53;  Mismatches   85;  Indels   40;  Gaps    3;

Qy        438 QAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQV------------------VQAKKFS 478
              |  : |  ||::  :|| |: |  ::  :   :                   :  |:
Db        217 QIRRKLRIEKRQQENLLLSVLPAHISMGMKLAIIERLKEHGDRRCMPDNNFHSLYVKRHQ 276

Qy        479 NVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL 538
              ||:|::|::|||||| : | ||| ::: :|| |: :|||     :  ::: :|| |    ||
Db        277 NVSILYADIVGFTQLASDCSPKELVVVLNELFGKFDQIAKANECMRIKILGDCYYCVSGL 336

Qy        539 HKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFG 598
                 |||     | | | :   :|    |  | ||:|:|||:|  ||:|::  :| ::
Db        337 PVSLPTHARNCVKMGLDMCQAIKQVREATGVDINMRVGIHSGNVLCGVIGLRKWQYDVWS 396

Qy        599 NNVTLANKFESCSVPRKINVSPTTYRLL------KDCPG--------------FVFTPR 637
              ::|:|||: |:  || :::::  | : |      :|  |                |  ||
Db        397 HDVSLANRMEAAGVPGRVHITEATLKHLDKAYEVEDGHGQQRDPYLKEMNIRTYLVIDPR 456

Qy        638 SREELPPN 645
              |::  ||:
Db        457 SQQPPPPS 464


RESULT 7
US-09-925-297-811
; Sequence 811, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA105
;  CURRENT APPLICATION NUMBER: US/09/925,297
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05989
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 928
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 811
;   LENGTH: 855
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (479)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (829)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-811

  Query Match             8.4%;  Score 303.5;  DB 10;  Length 855;
  Best Local Similarity   29.9%;  Pred. No. 1.3e-17;
  Matches   82;  Conservative   52;  Mismatches   93;  Indels   47;  Gaps    8;

Qy        433 KATLE-----QAHQALEEEKKKTVDLLCSIFPCEVAQQL-------------WQGQVVQA 474
              || ||     :    |||: ::  :|: || |  || ::              |   :
Db          7 KAFLEARQSLEVKMNLEEQSQQQENLMLSILPKHVADEMLKDMKKDESQKDQQQFNTMYM 66

Qy        475 KKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCV 534
              :  ||::||:|||||| : | ||  ::: :|| |: |||:   :    ::: :|| |
Db         67 YRHENVSILFADIVGFTQLSSACSAQELVKLLNELFARFDKLAAKYHQLRIKILGDCYYC 126

Qy        535 AGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRY 594
                ||    : |||   || | |:|    |       : ||:|:|:|:|  ||:| |  :|
Db        127 ICGLPDYREDHAVCSILMGLAMVEAISYVREKTKTGVDMRVGVHTGTVLGGVLGQKRWQY 186

Qy        595 CLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG-- 652
              ::  :||:||| |:  :| ::::| :|     ||            |   |  | ||
Db        187 DVWSTDVTVANKMEAGGIPGRVHISQST----MDC-----------LKGEFDVE-PGDG 229

Qy        653 --ICHFLDAYQQGTN------SKPCFQKKDVEDG 678
                  | :|:  ::|       |||  :|   ::|
Db        230 GSRCDYLE--EKGIETYLIIASKPEVKKTATQNG 261


RESULT 8
US-09-750-240-4
; Sequence 4, Application US/09750240
; Patent No. US20020103147A1
; GENERAL INFORMATION:
;  APPLICANT: Hammon, H. K.
;  APPLICANT:  Insel, P. A.
;  APPLICANT:  Ping, P.
;  APPLICANT:  Post, S. R.
;  APPLICANT:  Gao, M.
;  TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
;  TITLE OF INVENTION:  FAILURE
;  FILE REFERENCE: 220002056723
;  CURRENT APPLICATION NUMBER: US/09/750,240
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 09/472,667
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: US 09/008,097
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: US 08/924,757
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: US 60/048,933
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: US 08/708,661
;  PRIOR FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 604
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-750-240-4

  Query Match             7.9%;  Score 282.5;  DB 10;  Length 604;
  Best Local Similarity   31.8%;  Pred. No. 5.3e-16;
  Matches   63;  Conservative   44;  Mismatches   82;  Indels    9;  Gaps    1;

Qy        438 QAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG---------QVVQAKKFSNVTMLFSDIV 488
              ||   |: | ::   || |: |  || :: :            :  :|  ||::||:||
Db        115 QARLHLQHENRQQERLLLSVLPQHVAMEMKEDINTKKEDMFHKIYIQKHDNVSILFADIE 174

Qy        489 GFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQ 548
              |||:: |||:  :::  || |: |||:   |    ::: :|| |    || :    ||
Db        175 GFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEARADHAHC 234

Qy        549 IALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFE 608
                | : |:|    |    |  : ||:|:||| |  ||:|::  :: :: |:|||||  |
Db        235 CVEMGVDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHME 294

Qy        609 SCSVPRKINVSPTTYRLL 626
              : |   :|:::  | : |
Db        295 AGSRAGRIHITRATLQYL 312


RESULT 9
US-09-750-240-6
; Sequence 6, Application US/09750240
; Patent No. US20020103147A1
; GENERAL INFORMATION:
;  APPLICANT: Hammon, H. K.
;  APPLICANT:  Insel, P. A.
;  APPLICANT:  Ping, P.
;  APPLICANT:  Post, S. R.
;  APPLICANT:  Gao, M.
;  TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
;  TITLE OF INVENTION:  FAILURE
;  FILE REFERENCE: 220002056723
;  CURRENT APPLICATION NUMBER: US/09/750,240
;  CURRENT FILING DATE:  2001-10-12
```

```
;  PRIOR APPLICATION NUMBER: US 09/472,667
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: US 09/008,097
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: US 08/924,757
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: US 60/048,933
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: US 08/708,661
;  PRIOR FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 1167
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-750-240-6

  Query Match             7.9%;  Score 282.5;  DB 10;  Length 1167;
  Best Local Similarity   31.8%;  Pred. No. 1.4e-15;
  Matches   63;  Conservative   44;  Mismatches   82;  Indels    9;  Gaps    1;

Qy        438 QAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG---------QVVQAKKFSNVTMLFSDIV 488
              ||   |: | ::   || |: |  || :: :           :  :|  ||::||:||
Db        326 QARLHLQHENRQQERLLLSVLPQHVAMEMKEDINTKKEDMFHKIYIQKHDNVSILFADIE 385

Qy        489 GFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQ 548
              |||:: |||:  :::  || |: |||:   |    ::: :|| |    || :    ||
Db        386 GFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEARADHAHC 445

Qy        549 IALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFE 608
              | : |:|    |    |  : ||:|:||| |  ||:|::  :: :: |:|||||  |
Db        446 CVEMGVDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHME 505

Qy        609 SCSVPRKINVSPTTYRLL 626
              : |   :|:::  | : |
Db        506 AGSRAGRIHITRATLQYL 523


RESULT 10
US-10-175-158-2
; Sequence 2, Application US/10175158
; Publication No. US20030008371A1
; GENERAL INFORMATION:
;  APPLICANT: Tomlinson, James
;  APPLICANT:  Cor Therapeutics, Inc.
;  TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE: 44481-5027-01-US
;  CURRENT APPLICATION NUMBER: US/10/175,158
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US/09/473,716
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: PCT/US98/13540
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/070,901
;  PRIOR FILING DATE: 1997-07-01
;  PRIOR APPLICATION NUMBER: 08/886,362
;  PRIOR FILING DATE: 1997-07-01
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: human type V adenylyl cyclase
US-10-175-158-2

  Query Match             7.8%;  Score 280.5;  DB 9;  Length 1261;
  Best Local Similarity   26.4%;  Pred. No. 2.4e-15;
  Matches   83;  Conservative   52;  Mismatches  108;  Indels   71;  Gaps    8;

Qy        438 QAHQALEEEKKKTVDLLCSIFPCEVAQQL----------WQGQVVQAKKFSNVTMLFSDI 487
              ||    : | ::   || |: |  || ::              :  :|  ||::||:||

Db        416 QARLHSQRENQQQERLLLSVLPRHVAMEMKADINAKQEDMMFHKIYIQKHDNVSILFADI 475

Qy        488 VGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAV 547
              |||:: |||:  :::  || |: |||:   |    ::: :|| |    || :    ||
Db        476 EGFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEARADHAH 535

Qy        548 QIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKF 607
              | : |:|    |    |  : ||:|:||| |  ||:|::  :: :: |:|||||
Db        536 CCVEMGMDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHM 595

Qy        608 ESCSVPRKINVSPTT-------YRLLKDCPG-------------FVF---TPRSREEL-- 642
              |:     :|:::  |       | :   | |             |:    | : :||
Db        596 EAGGKAGRIHITKATLNYLNGDYEVEPGCGGERNAYLKEHSIETFLILRCTQKRKEEKAM 655

Qy        643 --------------PPNFPSEIPGICHFLDAYQQGTNSKPCFQKK----DVEDGNA--- 680
                           ||:: :| |   |       | |      |:    | :| ||
Db        656 IAKMNRQRTNSIGHNPPHWGAERPFYNHL------GGNQVSKEMKRMGFEDPKDKNAQES 709

Qy        681 --------NFLGKA 686
                       |||:|
Db        710 ANPEDEVDEFLGRA 723


RESULT 11
US-10-201-000-2
; Sequence 2, Application US/10201000
; Publication No. US20020187540A1
; GENERAL INFORMATION:
;  APPLICANT: Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE: 44481-5028-01-US
;  CURRENT APPLICATION NUMBER: US/10/201,000
;  CURRENT FILING DATE:  2002-07-24
;  PRIOR APPLICATION NUMBER: US/09/474,076
;  PRIOR FILING DATE: 1999-12-12
;  PRIOR APPLICATION NUMBER: PCT/US98/13694
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/070,904
;  PRIOR FILING DATE: 1997-07-01
;  PRIOR APPLICATION NUMBER: 08/886,550
;  PRIOR FILING DATE: 1997-07-01
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1168
;   TYPE: PRT
;   ORGANISM: human type VI adenylyl cyclase
US-10-201-000-2

  Query Match             7.8%;  Score 279.5;  DB 9;  Length 1168;
  Best Local Similarity   30.4%;  Pred. No. 2.6e-15;
  Matches   63;  Conservative   47;  Mismatches   82;  Indels   15;  Gaps    2;

Qy        430 GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG----------QVVQAKKFSN 479
              | ::| |   |:  ::|:     || |: |  || :: :             :  :|  |
Db        323 GYIQARLHLQHENRQQER-----LLLSVLPQHVAMEMKEDINTKKEDMMFHKIYIQKHDN 377

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539
              |::||:|| |||:: |||:  :::  || |: |||:   |    ::: :|| |    ||
Db        378 VSILFADIEGFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLP 437

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
              :    ||     | : |:|    |    |  : ||:|:||| |  ||:|::  :: :: |
Db        438 EARADHAHCCVEMGVDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSN 497

Qy        600 NVTLANKFESCSVPRKINVSPTTYRLL 626
              :|||||  |:     :|:::  | : |
Db        498 DVTLANHMEAGGRAGRIHITRATLQYL 524


RESULT 12
```

```
US-09-750-240-11
; Sequence 11, Application US/09750240
; Patent No. US20020103147A1
; GENERAL INFORMATION:
;  APPLICANT: Hammon, H. K.
;  APPLICANT:  Insel, P. A.
;  APPLICANT:  Ping, P.
;  APPLICANT:  Post, S. R.
;  APPLICANT:  Gao, M.
;  TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
;  TITLE OF INVENTION:  FAILURE
;  FILE REFERENCE: 220002056723
;  CURRENT APPLICATION NUMBER: US/09/750,240
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 09/472,667
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: US 09/008,097
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: US 08/924,757
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: US 60/048,933
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: US 08/708,661
;  PRIOR FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 1168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-750-240-11

  Query Match             7.8%;  Score 279.5;  DB 10;  Length 1168;
  Best Local Similarity   30.4%;  Pred. No. 2.6e-15;
  Matches   63;  Conservative   47;  Mismatches   82;  Indels   15;  Gaps    2;

Qy        430 GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG----------QVVQAKKFSN 479
              | ::| |   |:  ::|:     || |: |  || :: :             :  :|  |
Db        323 GYIQARLHLQHENRQQER-----LLLSVLPQHVAMEMKEDINTKKEDMMFHKIYIQKHDN 377

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539
              |::||:|| |||:: |||:  :::  || |: |||:   |    ::: :|| |    ||
Db        378 VSILFADIEGFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLP 437

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
              :    ||     | : |:|    |    | : ||:|:||| |  ||:|::  :: :: |
Db        438 EARADHAHCCVEMGVDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSN 497

Qy        600 NVTLANKFESCSVPRKINVSPTTYRLL 626
              :|||||  |:    :|:::  | : |
Db        498 DVTLANHMEAGGRAGRIHITRATLQYL 524


RESULT 13
US-09-750-240-13
; Sequence 13, Application US/09750240
; Patent No. US20020103147A1
; GENERAL INFORMATION:
;  APPLICANT: Hammon, H. K.
;  APPLICANT:  Insel, P. A.
;  APPLICANT:  Ping, P.
;  APPLICANT:  Post, S. R.
;  APPLICANT:  Gao, M.
;  TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
;  TITLE OF INVENTION:  FAILURE
;  FILE REFERENCE: 220002056723
;  CURRENT APPLICATION NUMBER: US/09/750,240
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 09/472,667
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: US 09/008,097
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: US 08/924,757
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: US 60/048,933
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: US 08/708,661
;  PRIOR FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 1167
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Modified AC-VI
US-09-750-240-13

  Query Match             7.7%;  Score 278;  DB 10;  Length 1167;
  Best Local Similarity   30.9%;  Pred. No. 3.5e-15;
  Matches   64;  Conservative   47;  Mismatches   80;  Indels   16;  Gaps    3;

Qy        430 GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQG----------QVVQAKKFSN 479
              | ::| |   |:  ::|:     || |: |  || :: :             :  :|  |
Db        323 GYIQARLHLQHENRQQER-----LLLSVLPQHVAMEMKEDINTKKEDMMFHKIYIQKHDN 377

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539
              |::||:|| |||:: |||:  :::  || |: |||:   |    ::: :|| |    ||
Db        378 VSILFADIEGFTSLASQCTAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLP 437

Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
              :    ||     | : |:|    |    | : ||:|:||| |  ||:|::  :: :: |
Db        438 EARADHAHCCVEMGVDMIEAISLVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSN 497

Qy        600 NVTLANKFESCSVPRKINVSPTTYRLL 626
              :|||||  |:    :|:::  | : |
Db        498 DVTLANHMEA-GGGRRIHITRATLQYL 523


RESULT 14
US-10-071-223-2
; Sequence 2, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
;  APPLICANT: Storm, Daniel R.
;  APPLICANT:  Hacker, Beth
;  APPLICANT:  Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  APPLICANT:  University of Washington
;  TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE: 44481-5029-02-US
;  CURRENT APPLICATION NUMBER: US/10/071,223
;  CURRENT FILING DATE:  2002-02-11
;  PRIOR APPLICATION NUMBER: 09/473,717
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: PCT/US98/13541
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/098,559
;  PRIOR FILING DATE: 1997-07-01
;  PRIOR APPLICATION NUMBER: 08/886,440
;  PRIOR FILING DATE: 1997-07-01
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1294
;   TYPE: PRT
;   ORGANISM: human type IX adenylyl cyclase
US-10-071-223-2

  Query Match             7.3%;  Score 263.5;  DB 12;  Length 1294;
  Best Local Similarity   29.9%;  Pred. No. 7.7e-14;
  Matches   69;  Conservative   58;  Mismatches   89;  Indels   15;  Gaps    6;

Qy        402 DIPIHNALRDVVLIGEQAR--AQDGLKKRLGKLKATLEQ-AHQALEEEKKKTVDLLCSIF 458
              |: :  ||:: ::     |  | | :|:   : : :::: |  : :  |||:     ||
```

```
Db        322 DLEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKS-----SIQ 376

Qy        459 PCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCG 518
              :| : :: | ::      |::||:|||||| : :  |   :: :|| |: |||: |
Db        377 KAPIAFRPFKMQQIE-----EVSILFADIVGFTKMSANKSAHALVGLLNDLFGRFDRLCE 431

Qy        519 ELDVYKVETIGDA-YCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGL 577
              |    |: |:||  |||||     :| ||     | | |::  ::      | : ||:|:
Db        432 ETKCEKISTLGDCYYCVAGCPEPRAD-HAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGV 490

Qy        578 HSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKD 628
              |:|:|  |::|::  :: :: |:| |||  |   |  |:::|  | : | |
Db        491 HTGTVLCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYLDD 541


RESULT 15
US-09-751-100B-2
; Sequence 2, Application US/09751100B
; Patent No. US20020142436A1
; GENERAL INFORMATION:
;  APPLICANT: Medical Research Council
;  TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor
;  FILE REFERENCE: P27948A
;  CURRENT APPLICATION NUMBER: US/09/751,100B
;  CURRENT FILING DATE:  2000-12-28
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: Mouse
US-09-751-100B-2

  Query Match             7.3%;  Score 260.5;  DB 10;  Length 1353;
  Best Local Similarity   29.4%;  Pred. No. 1.5e-13;
  Matches   68;  Conservative   59;  Mismatches   89;  Indels   15;  Gaps    6;

Qy        402 DIPIHNALRDVVLIGEQAR--AQDGLKKRLGKLKATLEQ-AHQALEEEKKKTVDLLCSIF 458
              |: :  ||:: ::     |  | | :|:  : : :::: |  : :  |||:     ||
Db        322 DLEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKS-----SIQ 376

Qy        459 PCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCG 518
              :| : :: | ::      |::||:|||||| : :  |   :: :|| |: |||: |
Db        377 KAPIAFRPFKMQQIE-----EVSILFADIVGFTKMSANKSAHALVGLLNDLFGRFDRLCE 431

Qy        519 ELDVYKVETIGDA-YCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGL 577
              :    |: |:||  |||||     :| ||     | | |::  ::      | : ||:|:
Db        432 QTKCEKISTLGDCYYCVAGCPEPRAD-HAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGV 490

Qy        578 HSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKD 628
              |:|:|  |::|::  :: :: |:| |||  |   |  |:::|  | : | |
Db        491 HTGTVLCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYLDD 541


Search completed: June 27, 2003, 13:17:11
Job time : 37.3713 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 12:42:07 ; Search time 32.6287 Seconds
(without alignments)
2527.899 Million cell updates/sec

Title: US-09-762-767A-4
Perfect score: 3231
Sequence: 1 MYGFVNHALELLVIRNYGPE.....QVWFLSRKNTGTEETKQDDD 619

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3231	100.0	619	21	AAY51608	Human soluble guan
2	1883	58.3	787	22	ABB58149	Drosophila melanog
3	929.5	28.8	690	22	ABB60360	Drosophila melanog
4	789.5	24.4	690	21	AAY70475	Human cyclic nucle
5	789.5	24.4	690	21	AAY51607	Human soluble guan
6	789.5	24.4	690	23	ABG61940	Prostate cancer-as
7	730.5	22.6	717	23	ABG61878	Prostate cancer-as
8	726	22.5	676	22	ABB58042	Drosophila melanog
9	704.5	21.8	669	22	ABB69903	Drosophila melanog
10	685	21.2	649	22	ABB69902	Drosophila melanog
11	599	18.5	119	21	AAG00754	Human secreted pro
12	481.5	14.9	1061	23	AAU11280	Human atrionatriur
13	481.5	14.9	1061	23	AAU11281	Human atrionatriur
14	481.5	14.9	1075	22	ABB11783	Human ANP-A recept
15	468	14.5	1029	14	AAR38862	GC-A. Rattus ratt
16	458	14.2	1076	22	ABB71206	Drosophila melanog
17	455.5	14.1	1047	12	AAR10867	NPRB(Pro655, Glu65
18	454.5	14.1	1025	14	AAR38863	GC-B. Rattus ratt
19	453.5	14.0	1047	12	AAR10399	Human Natriuretic
20	451	14.0	1525	22	ABB71674	Drosophila melanog
21	443	13.7	1076	22	ABB59684	Drosophila melanog
22	429.5	13.3	566	22	ABB61743	Drosophila melanog
23	429.5	13.3	975	22	ABB64661	Drosophila melanog
24	407.5	12.6	1073	18	AAW32063	Human ST receptor
25	407.5	12.6	1073	19	AAW37371	Human ST receptor
26	407.5	12.6	1073	23	AAU08788	Human guanylin cyc
27	406.5	12.6	1172	22	ABB72034	Drosophila melanog
28	404	12.5	1075	14	AAR38861	GC-C. Rattus ratt
29	399.5	12.4	949	22	AAB83941	Amino acid sequenc
30	397.5	12.3	1115	22	AAB83940	Amino acid sequenc
31	390.5	12.1	1151	22	AAB83939	Amino acid sequenc
32	389	12.0	887	22	ABG03067	Novel human diagno
33	355.5	11.0	273	22	AAG67400	Partial amino acid
34	344.5	10.7	1253	20	AAW88524	Adenyl cyclase-1 (
35	311	9.6	1119	23	AAE17133	Human adenylyl and
36	309.5	9.6	1134	21	AAB02005	Adenylyl cyclase t
37	309	9.6	533	21	AAB02013	Adenylyl cyclase C
38	308.5	9.5	1248	21	AAB02012	Type VIII adenylyl
39	299.5	9.3	2248	22	ABB64494	Drosophila melanog
40	295	9.1	1597	22	ABB60567	Drosophila melanog
41	292	9.0	1167	22	ABB58194	Drosophila melanog
42	291.5	9.0	1294	20	AAW30601	Human type IX aden
43	291.5	9.0	1305	20	AAW88525	Adenyl cyclase typ
44	291.5	9.0	1353	17	AAR99251	Murine adenylate c
45	288	8.9	1098	21	AAB02009	Type V adenylyl cy

ALIGNMENTS

RESULT 1
AAY51608
ID AAY51608 standard; Protein; 619 AA.
XX
AC AAY51608;
XX
DT 26-MAY-2000 (first entry)
XX
DE Human soluble guanylylcyclase beta1 protein.
XX
KW Human; guanylylcyclase alpha1; hsGCalpha1; hsGCbeta1; soluble;
KW guanylylcyclase beta1; antiarteriosclerotic; vasotropic; hypotensive;
KW gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;
KW peripheral arterial occlusive disease; arterial hypertension.
XX
OS Homo sapiens.
XX
PN DE19837015-A1.
XX
PD 24-FEB-2000.
XX
PF 14-AUG-1998; 98DE-1037015.
XX
PR 14-AUG-1998; 98DE-1037015.
XX
PA (VASO-) VASOPHARM BIOTECH GMBH & CO KG.
XX
PI Schmidt H, Zabel U, Poller W;
XX
DR WPI; 2000-184044/17.
DR N-PSDB; AAZ88939.
XX
PT New human soluble guanylate cyclase alpha1/beta1 and the nucleic acid

PT encoding the subunits, useful for producing diagnostic antibodies, and
PT for somatic gene therapy of arteriosclerosis -
XX
PS Example 1; Page 16-17; 44pp; German.
XX
CC This invention describes novel purified human soluble guanylate cyclase
CC alpha1/beta1 (hsGCalpha1/beta1). The products of the invention have
CC antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid
CC sequences encoding the alpha1 and/or beta1 subunit are useful for somatic
CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),
CC peripheral arterial occlusive disease and arterial hypertension.
CC Antibodies to hsGCalpha1/beta1 can be used for diagnosis of aberrant
CC hsGCalpha1/beta1 expression in human tissues. This sequence represents
CC the human soluble guanylylcyclase beta1 subunit described in the method
CC of the invention.
XX
SQ Sequence 619 AA;

Query Match 100.0%; Score 3231; DB 21; Length 619;
Best Local Similarity 100.0%; Pred. No. 5.4e-306;
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              |||||||||||||||||||
Db        601 VWFLSRKNTGTEETKQDDD 619
```

RESULT 2
ABB58149
ID ABB58149 standard; Protein; 787 AA.
XX
AC ABB58149;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 1239.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL02252.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 1239; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 787 AA;

Query Match 58.3%; Score 1883; DB 22; Length 787;
Best Local Similarity 49.9%; Pred. No. 3.3e-174;
Matches 371; Conservative 102; Mismatches 125; Indels 146; Gaps 7;

```
Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||:||||||::::| |:|| |||:| :  ||||||| ||||  ||:|: || ::||
Db          1 MYGFVNYALELLVLKHFGEEIWEKIKKKAMVSMEGQFLVRQIYDDEITYNLIGAAVEILN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              : | :||::||| || |||:|||| ||:|||:  |:||||||||||| |:|||||||||
Db         61 IPADDILELFGKTFFEFCQDSGYDKILQVLGATPRDFLQNLDALHDHLGTLYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNE-------- 172
              |||  ||   |:|||||| ||: |||||:| || ::|| |::: ::::: |
Db        121 RCT--EKDGELLLHYYSERPGLEHIVIGIVKAVASKLHGVEVEIDIVKRKGEPIDEAEKE 178

Qy        173 ------------------------------------------------------------ 172

Db        179 RAIARENQQLLEDAVATTTTGSATVVLAPSSDAERNNNHNGSNGSNNNGMANNGNTVNVN 238

Qy        173 ----------------------------EC---------------------DHTQFLIE 182
                                          :|                     || ||||
Db        239 NNNDGQQIASETDPSIALSTCPIAQDSFDCDGDKEQKCLRLLKNKSDDIERYDHVQFLIR 298

Qy        183 E--------KESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGN 234
              |          ::|:::  :|::       :   ||| |||| |||||::||| : : | |
Db        299 EINVAAKSQVDAKKDEVPDDMEFL----CEAPLISPATFCKVFPFHLMFDRQMKIVQAGK 354
```

```
Qy        235 AIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECED 294
Db        355 AVSRVIPRVAEENCSLIEVVEAIRPHLQLNFENILSHINTIYVLQTRQGAMS-------- 406

Qy        295 ELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVL 354
Db        407 --SRHEQRFLRLKGQMMYIPETDRILFQCYPSVMNLDDLTKKGLYISDVPLHDAARDLVL 464

Qy        355 LGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPA 414
Db        465 LSEKFEAEYKLTKNLEMLTDKLQQTFRDLESEKQKTDRLLYSVLPKSVANELRHQRPVPP 524

Qy        415 KRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVET 474
Db        525 KRYDSVTLMFSGIVGFGQYCAANTDPDGAMKIVKMLNELYTVFDALTDSKRNLNVYKVET 584

Qy        475 VGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIG 534
Db        585 VGDKYMAVSGLPDHCEDHAKCMARVALDMMDMAKNVKMGSNPVQITIGIHSGEVVTGVIG 644

Qy        535 QRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKG 594
Db        645 NRVPRYCLFGNTVNLTSRTETTGVPGRINVSEETYRLLCMAINQDDSFHLEYRGPVIMKG 704

Qy        595 KKEPMQVWFLSRKNT----GTEET 614
Db        705 KPTPMDCWFLTRATSSILGGTSST 728
```

RESULT 3
ABB60360
ID ABB60360 standard; Protein; 690 AA.
XX
AC ABB60360;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 7872.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL04463.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 7872; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 690 AA;

Query Match 28.8%; Score 929.5; DB 22; Length 690;
Best Local Similarity 34.8%; Pred. No. 3.3e-81;
Matches 218; Conservative 129; Mismatches 231; Indels 49; Gaps 15;

```
Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
Db          1 MYGLLLENLSEYIKSVYGEEKWEDIRRQAGIDSP-SFSVHQVYPENLLQKLAKKAQQVLG 59

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
Db         60 VSERDFMDQMGVYFVGFVGQYGYDRVLSVLGRHMRDFLNGLDNLHEYLKFSYPRMRAPSF 119

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQR---------- 170
Db        120 ICENETK-QGLTLHYRSKRRGFVYYTMGQIREVARYFYHKEMHIELVREEILFDTVHVTF 178

Qy        171 NEECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVT 230
Db        179 QLTFDNRAFTLASLAMTREE------------KHLPISAHVLFEIFPFCMVFGADMVVR 225

Qy        231 QCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKL 290
Db        226 SIGNSLMVILPELL--GKKITAWFDLVRPLIAFKFQTILNRTNNIFELVTVDPVTERFDV 283

Qy        291 ECEDEL---TGTE-ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346
Db        284 QNEDLLQHEDGSEPEKSLRLKGQMVYMENWRMIMFLGTPVMPDLTSLITTGLYINDLSMH 343

Qy        347 DATRDLVLLGEQFREEYKLTQELE-ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANE 405
Db        344 DFSRDLMLAGTQQSVELKLALDQEQQKSKKLEESMRLLDEEMRRTDELLYQMIPKQVADR 403

Qy        406 LRH-KRPV-PAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDS 463
Db        404 LRRGENPIDTCEMFDSVSILFSDIVTFTEICSRIT----PMEVVSMLNAMYSIFDKLTER 459

Qy        464 RKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQ--VDGESVQITI 521
Db        460 NS---VYKVETIGDAYMVVAGAPDKDANHAERVCDMALDMVDAITDLKDPSTGQHLRIRV 516

Qy        522 GIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQ 581
Db        517 GVHSGAVVAGIVGLKMPRYCLFGDTVNTASRMESTSIAMKVHISEST-KVLIG-----PN 570

Qy        582 FHLEHRGPVSMKGKKEPMQVWFLSRKN 608
Db        571 YKIIERGEIDVKGKGTMGTYWLEEREN 597
```

RESULT 4
AAY70475
ID AAY70475 standard; Protein; 690 AA.
XX
AC AAY70475;
XX
DT 04-JUL-2000 (first entry)
XX
DE Human cyclic nucleotide-associated protein-3 (CNAP-3).
XX
KW Cyclic nucleotide-associated protein-3; CNAP-3; human; cytostatic;
KW anti-arteriosclerotic; hepatotropic; anti-leukaemic; anti-inflammatory;
KW immunomodulatory; anti-asthmatic; anti-anaemic; anti-diabetic; diagnosis;
KW anti-sclerotic; dermatological; neuroprotective; anti-epileptic; cancer;
KW anti-Alzheimer's; anti-Parkinsonian; cerebroprotective; ophthalmological;
KW anti-infertility; anti-allergic; vasotropic; immunosuppressive;
KW hypotensive; gene therapy; prevention; treatment; arteriosclerosis;

```
KW   cell proliferative disorder; autoimmune/inflammatory; diabetes mellitus;
KW   neurological; vision; reproductive; smooth muscle.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   65
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   211
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   361
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   450
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   478
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   29
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   73
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   117
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   135
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   158
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   229
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   511
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   527
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   638
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   687
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   39
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   64
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   105
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   261
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   328
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   356
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   360
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   394
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   622
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   635
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   27
FT                   /note= "N-glycosylated"
FT   Modified-site   332
FT                   /note= "N-glycosylated"
FT   Modified-site   479
FT                   /note= "N-glycosylated"
FT   Modified-site   600
FT                   /note= "N-glycosylated"
FT   Region          472..661
FT                   /note= "Guanylate cyclase signature sequence
FT                   identified by PFAM analysis; CNAP-3 shares 89%
FT                   identity with human soluble guanylate cyclase
FT                   large subunit"
XX
PN   WO200014248-A1.
XX
PD   16-MAR-2000.
XX
PF   03-SEP-1999;   99WO-US20287.
XX
PR   04-SEP-1998;   98US-0148904.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Yue H,  Guegler KJ,  Corley NC,  Patterson C,  Tang YT;
XX
DR   WPI; 2000-256994/22.
DR   N-PSDB; AAZ51684.
XX
PT   Isolated cyclic nucleotide associated proteins useful for preventing,
PT   diagnosing and treating cell proliferative, autoimmune/inflammatory,
PT   neurological, vision, reproductive and smooth muscle disorders -
XX
PS   Disclosure; Page 67-69; 78pp; English.
XX
CC   The present sequence is a human cyclic nucleotide
CC   associated protein-3 (CNAP-3), identified in Incyte clone 159278,
CC   that is isolated from ADENINB01 cDNA library. It is expressed in
CC   nervous, reproductive, cardiovascular and developmental tissues.
CC   CNAP sequences may be used for prevention, treatment and diagnosis of
CC   diseases associated with altered CNAP expression such as, cell
CC   proliferative disorders (e.g. arteriosclerosis, cirrhosis, leukaemia,
CC   lymphoma and cancer of the breast, prostate, lung and brain), autoimmune/
CC   inflammatory disorders (e.g. asthma, anaemia, diabetes mellitus, multiple
CC   sclerosis and psoriasis), neurological disorders (e.g. epilepsy,
CC   Alzheimer's/Parkinson's disease and strokes), vision disorders (e.g.
CC   conjunctivitis, glaucoma, cataracts and retinitis pigmentosa),
CC   reproductive disorders (e.g. infertility, uterine fibroids, ectopic
CC   pregnancies and impotence) and smooth muscle disorders (e.g. angina,
CC   anaphylactic shock, Kearns-Sayre syndrome and hypertension). The
CC   coding sequence can be used for gene therapy.
XX
SQ   Sequence   690 AA;

  Query Match             24.4%;  Score 789.5;  DB 21;  Length 690;
  Best Local Similarity   35.1%;  Pred. No. 1.5e-67;
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88
              ::|  : | |  :: :     |||   |      |:: :  |:  |  | |   : ||
Db        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153

Qy         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143
              |:|  :::|| :   |     |   |      :   | | | |    | ::|:  :
Db        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212

Qy        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192
               |: ||||  |  :: ||::: ::      | |           : :  | :|
Db        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265

Qy        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250
                     :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320

Qy        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310
                | :: | |: :| ||:: :|  ||:| :     |:|               : |||||
Db        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370

Qy        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370
              ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430

Qy        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430
                |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490

Qy        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490
               | ||: :     :::: :|| ||||||          ||||||:|| |    || :
Db        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543
```

```
Qy        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549
              ||  |  :|| |||:: :|    || ::: ||:|:| |  ||:| :|||||||| | |
Db        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603

Qy        550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588
              :: |:     |||||  ||| |           | |   | |  |  |
Db        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652


RESULT 5
AAY51607
ID   AAY51607 standard; Protein; 690 AA.
XX
AC   AAY51607;
XX
DT   26-MAY-2000  (first entry)
XX
DE   Human soluble guanylylcyclase alpha1 protein.
XX
KW   Human; guanylylcyclase alpha1; hsGCalpha1; hsGCbeta1; soluble;
KW   guanylylcyclase beta1; antiarteriosclerotic; vasotropic; hypotensive;
KW   gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;
KW   peripheral arterial occlusive disease; arterial hypertension.
XX
OS   Homo sapiens.
XX
PN   DE19837015-A1.
XX
PD   24-FEB-2000.
XX
PF   14-AUG-1998;   98DE-1037015.
XX
PR   14-AUG-1998;   98DE-1037015.
XX
PA   (VASO-) VASOPHARM BIOTECH GMBH & CO KG.
XX
PI   Schmidt H,  Zabel U,  Poller W;
XX
DR   WPI; 2000-184044/17.
DR   N-PSDB; AAZ88938.
XX
PT   New human soluble guanylate cyclase alpha1/beta1 and the nucleic acid
PT   encoding the subunits, useful for producing diagnostic antibodies, and
PT   for somatic gene therapy of arteriosclerosis -
XX
PS   Example 1; Page 13-15; 44pp; German.
XX
CC   This invention describes novel purified human soluble guanylate cyclase
CC   alpha1/beta1 (hsGCalpha1/beta1). The products of the invention have
CC   antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid
CC   sequences encoding the alpha1 and/or beta1 subunit are useful for somatic
CC   gene therapy of arteriosclerosis and restenosis, ischemia (infarct),
CC   peripheral arterial occlusive disease and arterial hypertension.
CC   Antibodies to hsGCalpha1/beta1 can be used for diagnosis of aberrant
CC   hsGCalpha1/beta1 expression in human tissues. This sequence represents
CC   the human soluble guanylylcyclase alpha1 subunit described in the method
CC   of the invention.
XX
SQ   Sequence   690 AA;

  Query Match             24.4%;  Score 789.5;  DB 21;  Length 690;
  Best Local Similarity   35.1%;  Pred. No. 1.5e-67;
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88
              :: |  : | |  :: :     |||   |      |:: :  |:  |  | |   : ||
Db        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153

Qy         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143
              |:|  :::|| :   |     |   |     :   | | | |    | ::|:  :
Db        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212

Qy        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192
              |: ||||  |  :: ||::: ::      | |            : :  | :|
Db        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265

Qy        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250
                    :  | | : | : ||| |||| :||:|: : | || | |:: :    | :
Db        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320

Qy        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310
              | :: | |: :| ||:: :|  ||:| :     |:|             : |||||
Db        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370

Qy        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370
              ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430

Qy        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430
              |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490

Qy        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490
              | ||: :     :::: :|| ||||||          ||||||:|| |    || :
Db        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543

Qy        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549
              ||  |  :|| |||:: :|    || ::: ||:|:| |  ||:| :|||||||| | |
Db        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603

Qy        550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588
              :: |:     |||||  ||| |           | |   | |  |  |
Db        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652


RESULT 6
ABG61940
ID   ABG61940 standard; Protein; 690 AA.
XX
AC   ABG61940;
XX
DT   15-AUG-2002  (first entry)
XX
DE   Prostate cancer-associated protein #141.
XX
KW   Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS   Mammalia.
XX
PN   WO200230268-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32045.
XX
PR   13-OCT-2000; 2000US-0687576.
PR   08-DEC-2000; 2000US-0733288.
PR   08-DEC-2000; 2000US-0733742.
PR   24-JAN-2001; 2001US-263957P.
PR   16-MAR-2001; 2001US-276791P.
PR   16-MAR-2001; 2001US-276888P.
PR   06-APR-2001; 2001US-281922P.
PR   24-APR-2001; 2001US-286214P.
PR   30-APR-2001; 2001US-0847046.
PR   04-MAY-2001; 2001US-288589P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Gish KC,  Mack DH,  Wilson KE,  Afar D,  Hevezi P;
XX
DR   WPI; 2002-471335/50.
DR   N-PSDB; ABK92259.
XX
PT   Detecting a prostate cancer-associated transcript in a cell in a
PT   patient, useful for diagnosing prostate cancer (PC) or screening
PT   modulators of PC, by determining if prostate cancer-associated genes
```

```
PT   are expressed in a prostate tissue -
XX
PS   Claim 27; Page 421-422; 436pp; English.
XX
CC   The present invention relates to methods of detecting a prostate
CC   cancer-associated transcript in a cell from a patient. The method
CC   comprises contacting a biological sample from the patient with
CC   prostate cancer-associated polynucleotides (designated PC genes) that
CC   selectively hybridise to a sequence that is at least 80% identical
CC   to them. The prostate cancer-associated polynucleotide sequences
CC   are differentially expressed in prostate tumour tissue or in
CC   prostate cancer and are derived from the tissues of various
CC   organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC   The methods of the invention are useful for diagnosing and treating
CC   prostate cancer in mammals. The prostate cancer-associated genes are
CC   useful for diagnosing or treating prostate cancer, as well as for
CC   identifying modulators of prostate cancer or agents that inhibit
CC   prostate cancer. The nucleic acid sequences are particularly useful
CC   in gene therapy, as a vaccine or in antisense applications.
CC   ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
SQ   Sequence   690 AA;

  Query Match             24.4%;  Score 789.5;  DB 23;  Length 690;
  Best Local Similarity   35.1%;  Pred. No. 1.5e-67;
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy          30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88
               ::  |  : | |  :: :     |||   |      |:: :  |:  |  | |   : ||
Db         101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153

Qy          89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143
               |:|  :::|| :   |     |   |     :   |  | | |    | ::|:  :
Db         154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212

Qy         144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ-----------FLIEEKESKEEDFY 192
               |: ||||  |  :: ||::: ::      | |             : :  | :|
Db         213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265

Qy         193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250
                     :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db         266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320

Qy         251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310
                 | :: | |: :| ||:: :|  ||:| :     |:|             : |||||
Db         321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370

Qy         311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370
               ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db         371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430

Qy         371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430
                 |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db         431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490

Qy         431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490
                 | ||: :     :::: :|| ||||||          ||||||:|| |    || :
Db         491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543

Qy         491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549
                 ||  |  :|| |||:: :|    || ::: ||:|:| |  ||:| :||||||||| | |
Db         544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603

Qy         550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588
               :: |:     |||||  ||| |           | |   | |  | |
Db         604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652


RESULT 7
ABG61878
ID   ABG61878 standard; Protein; 717 AA.
XX
AC   ABG61878;
XX
DT   15-AUG-2002  (first entry)
XX
DE   Prostate cancer-associated protein #79.
XX
KW   Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS   Mammalia.
XX
PN   WO200230268-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32045.
XX
PR   13-OCT-2000; 2000US-0687576.
PR   08-DEC-2000; 2000US-0733288.
PR   08-DEC-2000; 2000US-0733742.
PR   24-JAN-2001; 2001US-263957P.
PR   16-MAR-2001; 2001US-276791P.
PR   16-MAR-2001; 2001US-276888P.
PR   06-APR-2001; 2001US-281922P.
PR   24-APR-2001; 2001US-286214P.
PR   30-APR-2001; 2001US-0847046.
PR   04-MAY-2001; 2001US-288589P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Gish KC,  Mack DH,  Wilson KE,  Afar D,  Hevezi P;
XX
DR   WPI; 2002-471335/50.
DR   N-PSDB; ABK92193.
XX
PT   Detecting a prostate cancer-associated transcript in a cell in a
PT   patient, useful for diagnosing prostate cancer (PC) or screening
PT   modulators of PC, by determining if prostate cancer-associated genes
PT   are expressed in a prostate tissue -
XX
PS   Claim 27; Page 363; 436pp; English.
XX
CC   The present invention relates to methods of detecting a prostate
CC   cancer-associated transcript in a cell from a patient. The method
CC   comprises contacting a biological sample from the patient with
CC   prostate cancer-associated polynucleotides (designated PC genes) that
CC   selectively hybridise to a sequence that is at least 80% identical
CC   to them. The prostate cancer-associated polynucleotide sequences
CC   are differentially expressed in prostate tumour tissue or in
CC   prostate cancer and are derived from the tissues of various
CC   organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC   The methods of the invention are useful for diagnosing and treating
CC   prostate cancer in mammals. The prostate cancer-associated genes are
CC   useful for diagnosing or treating prostate cancer, as well as for
CC   identifying modulators of prostate cancer or agents that inhibit
CC   prostate cancer. The nucleic acid sequences are particularly useful
CC   in gene therapy, as a vaccine or in antisense applications.
CC   ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
SQ   Sequence   717 AA;

  Query Match             22.6%;  Score 730.5;  DB 23;  Length 717;
  Best Local Similarity   36.9%;  Pred. No. 9.3e-62;
  Matches  182;  Conservative   80;  Mismatches  178;  Indels   53;  Gaps   10;

Qy         119 SFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ 178
               |  | | |    | ::|:  :     |: ||||  |  :: ||::: ::      | |
Db         187 SILCLDKE-DDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLM----PPCFHND 241

Qy         179 -----------FLIEEKESKEEDFYEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRD 226
                         : :  | :|           :  | | : | : ||| |||| :||:|
Db         242 CSEFVNQPYLLYSVHMKSTKPS--------LSPSKPQSSLVIPTSLFCKTFPFHFMFDKD 293

Qy         227 LVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLD 286
               : : | || | |:: :           | :: | |: :| ||:: :|  ||:| :
```

```
Db        294 MTILQFGNGIRRLMNRRDFQGKPNFEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNS 353

Qy        287 VEKLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346
              |:|             : |||||||: |: :|||| || |  |:| | |||||||||:|
Db        354 VKK----------SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIH 403

Qy        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406
              :| ||:||:||| | :  | : |  |   |:   :|||:|||||  || |: |  || :|
Db        404 NALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQL 463

Qy        407 RHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKN 466
                 : | ||:: |||:||| |||| | ||: :     :::: :|| ||||||
Db        464 WQGQVVQAKKFSNVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFDQQCGELD- 518

Qy        467 PFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHT 525
                ||||||:    : : || :    ||  |  :|| |||:: :|    || ::: ||:|:
Db        519 --VYKVETIAMPIVWLGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHS 576

Qy        526 GEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL----------MSP 575
              | |  ||:| :|||||||||| | | :: |:     |||||  ||| |           |
Db        577 GSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSR 636

Qy        576 ENSDPQFHLEHRG 588
              |   | |  |  |
Db        637 EELPPNFPSEIPG 649


RESULT 8
ABB58042
ID   ABB58042 standard; Protein; 676 AA.
XX
AC   ABB58042;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 918.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02145.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 918; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   676 AA;

  Query Match             22.5%;  Score 726;  DB 22;  Length 676;
  Best Local Similarity   30.4%;  Pred. No. 2.3e-61;
  Matches  187;  Conservative  125;  Mismatches  235;  Indels   68;  Gaps   18;

Qy         22 WEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVL----NLNAGEILQMFGKMFFVF 77
              | :| |  :|| :        :     || :|   ::|      :| ||| : |:
Db         73 WPNIHK-LKLDPQ-------TFKSCANYDYLADIQELLLKMDEASASEILVLLGEELITC 124

Qy         78 CQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDAEKGKGLILHYYS 137
              |    :   | ||:::::||| :|| ::| |       :    | |      |:| :: : |
Db        125 CCTGIIERAFRCLGTDLQEFLGSLDGVYDVLKLQEEDVTDTGFVCA----GEGELI-FTS 179

Qy        138 EREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEE----------CDHTQ-FLIEEKES 186
              ||  :  :::| :| : : ::  ::::|:      :            |::|  |:    |
Db        180 ERPVIAWLLLGSLKALTRMLYKVDVNIKIEPVEGDARRYRYLFSLVKDNSQTMLMGRPTS 239

Qy        187 KEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL-PQLQP 245
                :   | : |   :   : :::  :||| ||:| | :  | : | |     ::  | :
Db        240 VSKTIPETVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLYKPYMAD 299

Qy        246 GNCSLLSVFSLVRPH-IDISFHGILSHINTVFV--LRSKEGLLDVEKLECEDELTGTEIS 302
              |   : |   ||  : : |  |:    | |:  | :  | :|   :
Db        300 FGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIG---------- 348

Qy        303 CLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREE 362
              | :|||:: ||::|:||: || :  || ||  ||::|||||||||||:::|:||| | :
Db        349 -LEIKGQMVHCPESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQARAQ 407

Qy        363 YKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTI 422
              | : :: : : ::    |:  |:||  :||: : |  :| :|     : || | :|||
Db        408 DGLRRRMDKIKNSIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYPDVTI 467

Qy        423 LFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPF--VYKVETVGDKYM 480
              ||| |||| : ||:         ::::|  ||  ||   |      |  ||||||:|| |
Db        468 LFSDIVGFTSICSRAT----PFMVISMLEGLYKDFDEFCD-----FFDVYKVETIGDAYC 518

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEVVTGVIGQRMPR 539
              |||    |: |  :  :|| |::   : :  ||| ::: ||:||| |: ||:|::|||
Db        519 VASGLHRASIYDAHKVAWMALKMIDACSKHITHDGEQIKMRIGLHTGTVLAGVVGRKMPR 578

Qy        540 YCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPM 599
              |||||::| : :: |:  |  |||||  |   |   |    :| |: | | |   |: |
Db        579 YCLFGHSVTIANKFESGSEALKINVSPTTKDWLTKHEGF--EFELQPRDP-SFLPKEFP- 634

Qy        600 QVWFLSRKNTGTEET 614
                    | |  ||
Db        635 --------NPGGTET 641


RESULT 9
ABB69903
ID   ABB69903 standard; Protein; 669 AA.
XX
AC   ABB69903;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 36501.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
```

```
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL14006.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 36501; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   669 AA;

  Query Match             21.8%;  Score 704.5;  DB 22;  Length 669;
  Best Local Similarity   28.8%;  Pred. No. 2.9e-59;
  Matches  197;  Conservative  123;  Mismatches  264;  Indels  101;  Gaps   18;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDE--EGQFLVRIIYDDSKTYDLVAAASKV 58
              ||| :  :::  : : || | |   :|  |: :     |    || |    |  ||
Db          1 MYGMLYESVQHYIQQEYGMETW---RKVCQIVDCKHQSFKTHQIYPDKLMPDFAAA---- 53

Qy         59 LNLNAGE----ILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPG 114
              |: : ||     :  ||: |  |    ||| ::|  |    :|||::| :|  :   ||
Db         54 LSASTGESFDFCMNFFGRCFVRFFSNFGYDKMIRSTGRYFCDFLQSIDNIHVQMRFTYPK 113

Qy        115 MRAPSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC 174
              |::|| : |: :   | :: | | | |:   :|| :  ||:: :| ::   |:: :|: |
Db        114 MKSPSMQLTNMD-DDGAVILYRSGRTGMSKYLIGQMTEVAKEFYGLDMTAYVLESQNDIC 172

Qy        175 DHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPY-------------TFCKAFPFH 220
              |   |:  |·      :    |:       |::                 | : |||
Db        173 GGTAGPIKLTEGPLTVIVKYRLDFDNRDYMAKRVNVIAHPSQLKMPSVDLNVFLELFPFT 232

Qy        221 IIFDRDLVVTQCGNAIYRVLPQLQPG---------------NC---------------- 248
              |: | |: :|  |  |'       ||              |
Db        233 IVLDHDMKITLAGEKIVETWILHNPGVNPKTFIGSHILERFKCRRPKDTQIQWETILQMR 292

Qy        249 SLLSVFSLVR---------PHIDISFHGI-----LSHINTVFVLRSKEGLLDVEKLEC-- 292
              ::|  | |:|           ::  |         |:    : :  :||   :  | |
Db        293 TVLFEFELIRTGHNRAAYDAALNFDFENFDEASSLNEAQAMALASAKEFSAENAKEEAAA 352

Qy        293 ----EDEL---TGTE-----ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYL 340
                  :||:   ||       :  : |||| |: : ||::|||| : |||:|   ||||
Db        353 AATSKDEIDPATGQRRHSVGLRSILLKGQMFYIKDVDSLIFLCSPLIENLDELHGIGLYL 412

Qy        341 SDIPLHDATRDLVLLGEQFREEYKLT-QELEILTDRLQLTLRALEDEKKKTDTLLYSVLP 399
              :|:  |  :|:||: | |   : ::  :: |  :| |: :|   :  |:: | ||||::|
Db        413 NDLNPHGLSRELVMAGWQHCSKLEIMFEKEEQRSDELEKSLELADSWKRQGDELLYSMIP 472

Qy        400 PSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDT 459
              :|  :|       : :: |:::|  ::  | : |   : : ||: |  || :   |
Db        473 RPIAERMRKSEEHVCQSFEEVSVIFIEVM--NIYDSGSNNIQDAMQAVTTLNKV---FSA 527

Qy        460 LTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQI 519
```

```
              | :   :|||||||||||  || ||| |:    ||   | ||| :|:      : |  | |
Db        528 LDEEIISPFVYKVETVGMVYMAVSGAPDVNPLHAEHACDLALRVMKKVKAHALPG--VAI 585

Qy        520 TIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSD 579
              :||::| || ||:| ::|||||||:|||  || |:: :   | :| ||     : :
Db        586 RVGINSGPVVAGVVGMKVPRYCLFGDTVNTASRMESSSDPWMIQLSNYT-----ALKVQK 640

Qy        580 PQFHLEHRGPVSMKGKKEPMQVWFL 604
                : :| || | :||| |    | |
Db        641 VGYKVEARGFVKVKGKGEMETYWLL 665


RESULT 10
ABB69902
ID   ABB69902 standard; Protein; 649 AA.
XX
AC   ABB69902;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 36498.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL14005.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 36498; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   649 AA;

  Query Match             21.2%;  Score 685;  DB 22;  Length 649;
  Best Local Similarity   28.5%;  Pred. No. 2.2e-57;
  Matches  193;  Conservative  122;  Mismatches  256;  Indels  106;  Gaps   18;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||| :  :::  |   || ::|  :       :   |    || |    |:  | |
Db          1 MYGMLYESVQHYVQEEYGVDIWRKVCHIIDC-KHNSFKTHQIYPDKLMPDIAEALSACTG 59

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
                |                   ||| ::|  |    :|||::| :|  :   || |::||
```

```
Db         60 DN-----------------FGYDKMIRSTGRYFCDFLQSIDNIHLIMRFTYPKMKSPSM 101

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHT--- 177
              : |: :   | :: | | | |:   :|| :  ||:: :| ||   ||: :|:    |
Db        102 QLTNMD-DNGAVILYRSSRTGMSKYLIGQMTEVAREFYGLEIKAYVIESQNDISGGTAGP 160

Qy        178 ------QFLIEEKESKEEDFYEDLDRFEENGTQESRISPYT-----FCKAFPFHIIFDRD 226
                   :  |   : | | : :       |::   |     |   |||  : : |
Db        161 IKLTDGPLTVIVKYRLDFDNREYMAKRVNTEAHPSQLKMPTVKLDVFLDLFPFTFVLNHD 220

Qy        227 LVVTQCGNAIYRVLPQLQPG--------------NC----------------SLLSVF 254
              : :|  |  |         ||              |                ::|  |
Db        221 MKITHAGEKIVETWIMHNPGANPKSFIGTHVMDLFQCRRPKDTTIDWDTLIQMRAVLFEF 280

Qy        255 SLVRP-HIDISFHGI-----------LSHINTVFVLRSKE----GLLDVEKLECEDEL- 296
               |:|  |   ::  :           |:   |: : :::|      :| ::   |||:
Db        281 ELIRTGHNRAAYDAVLNMDFENYDEMDLNEAQTMALAKAQEFSESHPVDDDESAREDEID 340

Qy        297 --TGTE-----ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDAT 349
                ||       :  : ||||| |: : ||::||||| : |||:|   ||||:|:  |  :
Db        341 PATGERRSSQGLRSILLKGQMFYIKDVDSLIFLCSPLIENLDELHGIGLYLNDLNPHGLS 400

Qy        350 RDLVLLGEQFREEYKLT-QELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRH 408
              |:||: | |   : ::  :: |  :| |: :|   :  |:: | ||||::|  :|  :|
Db        401 RELVMAGWQHCSKLEIMFEKEEQRSDELEKSLELADSWKRQGDELLYSMIPRPIAERMRL 460

Qy        409 KRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPF 468
                  : :: |:::|  ::  | :     | :|||: || || :   |  | :   :||
Db        461 SEEQVCQSFEEVSVIFLEVM--NVYDEGLNSIQGAMQTVNTLNKV---FSALDEEIISPF 515

Qy        469 VYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQV-DGESVQITIGIHTGE 527
              ||||||||  || ||| |:    ||   | ||| :|:   : :  |   | | :||::|
Db        516 VYKVETVGMVYMAVSGAPDVNPLHAEHACDLALRVMK---KFKAHDMGDVAIRVGINSGP 572

Qy        528 VVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHR 587
              || ||:||::|||||||:|||  || |:: :  || :|:||       :     : :| |
Db        573 VVAGVVGQKVPRYCLFGDTVNTASRMESSSDPWKIQLSKYT-----GDKVRQVGYKVESR 627

Qy        588 GPVSMKGKKEPMQVWFL 604
              | | :||| :    | |
Db        628 GTVQVKGKGDMETYWLL 644


RESULT 11
AAG00754
ID   AAG00754 standard; Protein; 119 AA.
XX
AC   AAG00754;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 4835.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC00760.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 13; SEQ ID 4835; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   119 AA;

  Query Match             18.5%;  Score 599;  DB 21;  Length 119;
  Best Local Similarity   98.3%;  Pred. No. 3.3e-50;
  Matches  117;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPS 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIXXGMRAPS 119


RESULT 12
AAU11280
ID   AAU11280 standard; Protein; 1061 AA.
XX
AC   AAU11280;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human atrionatriuretic peptide receptor A (NPR1) protein.
XX
KW   Human; natriuretic peptide receptor A/guanylate cyclase A; NPR1;
KW   atrionatriuretic peptide receptor A; haplotyping; cytostatic; genotyping;
KW   haplotype pair; single nucleotide polymorphism; gene therapy;
KW   drug screening; hypertension; hypotensive.
XX
OS   Homo sapiens.
XX
PN   WO200179231-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US12300.
XX
PR   14-APR-2000; 2000US-197330P.
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Bentivegna SC,  Choi JY,  Kliem SE,  Nandabalan K;
XX
DR   WPI; 2002-066340/09.
DR   N-PSDB; AAS16995.
XX
PT   Genotyping human natriuretic peptide receptor A/guanylate cyclase gene
PT   of an individual, involves determining identity of nucleotide pair at
PT   specific polymorphic sites for two copies of the gene  -
XX
PS   Claim 27; Fig 3; 96pp; English.
XX
```

```
CC   The invention relates to single nucleotide polymorphisms in the gene
CC   encoding the human natriuretic peptide receptor A/guanylate cyclase A
CC   (atrionatriuretic peptide receptor A) or NPR1 polypeptide. A method for
CC   haplotyping the NPR1 gene in an individual comprises identifying the
CC   nucleotide at one or more polymorphic sites and determining whether one
CC   of the copies of the gene is defined by one of the NPR1 haplotypes given
CC   in the specification or whether both copies are defined by a haplotype
CC   pair. This method is useful in genotyping, whereby all possible haplotype
CC   pairs can be assigned to specific genotypes. An association between a
CC   trait and a haplotype or haplotype pair of the NPR1 gene can be
CC   identified by comparing the frequency of the haplotype or haplotype pair
CC   in a population exhibiting the trait with the frequency of the haplotype
CC   or haplotype pair in a reference population, where a higher haplotype
CC   frequency in the trait population indicates the trait is associated with
CC   the haplotype or haplotype pair. NPR1 and its corresponding DNA are used
CC   for studying the expression and function of NPR1, for use in screening
CC   for candidate drugs to treat diseases related to NPR1 activity, such as
CC   hypertension. The sequences are also useful for studying the effect of
CC   variation on the biological activity of NPR1 as well as on the binding
CC   affinity of candidate drugs targeting NPR1. This sequence represents the
CC   human NPR1 polypeptide.
XX
SQ   Sequence   1061 AA;

  Query Match            14.9%;  Score 481.5;  DB 23;  Length 1061;
  Best Local Similarity  27.7%;  Pred. No. 3.6e-37;
  Matches  197;  Conservative   88;  Mismatches  227;  Indels  199;  Gaps   31;

Qy         33 EEGQFLVRIIYDDSKTYDLVAAASKVLN-------------------------LNAGEI 66
              | | | | : |: : : :|:| : : ||                          |:  |:
Db        416 ENGAFRVVLNYNGT-SQELVAVSGRKLNWPLGYPPPDIPKCGFDNEDPACNQDHLSTLEV 474

Qy         67 LQMFGKM---------FFVF----CQESGYDTILRVLGSNVREFLQNLDALHDHLATIYP 113
              | : | :         ||::     ::     : ||   :|        :|  ||
Db        475 LALVGSLSLLGILIVSFFIYRKMQLEKELASELWRVRWEDVEP-----SSLERHL----- 524

Qy        114 GMRAPSFRCTDAEKGK--GLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRN 171
                |:   | | : :|   | :|      ||   :     ||   :  |  : :| : ::
Db        525 --RSAGSRLTLSGRGSNYGSLL----TTEGQFQV---FAKTAYYK--GNLVAVKRVNRKR 573

Qy        172 EECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVT 230
              |   |: :: | :   :   | | ||    |    |   | :|       | : : :
Db        574 IEL--TRKVLFELKHMRDVQNEHLTRFVGACTDPPNICILTEYCPRGSLQDILENESITL 631

Qy        231 ------QCGNAIYRVLPQLQPGN-CSLLSVFSLVRPHIDISFHGILSHINTV----FVLR 279
                   | | : :  |  |  ||                 || |   | |    |||:
Db        632 DWMFRYSLTNDIVKGMLFLHNGAICS---------------HGNLKSSNCVVDGRFVLK 675

Qy        280 SKE-GLLDVEKLECEDELTGTEISCLRLKGQMIY---LPEADSILFLCSPSV--MNLDDL 333
                : ||     |: |              :|  :|   |  |  :|.: || |  .    |:
Db        676 ITDYGLESFRDLDPE------------QGHTVYAKKLWTAPELLRMASPPVRGSQAGDV 722

Qy        334 TRRGLYLSDIPL-----HDATRDL--------VLLGEQ--FREEYKLTQELEIL------ 372
                |: | :| |     |    ||        |  |||  ||     |   || |
Db        723 YSFGIILQEIALRSGVFHVEGLDLSPKEIIERVTRGEQPPFRPSLALQSHLEELGLLMQR 782

Qy        373 -----------TDRLQLTLR---------------------------------ALEDEK 387
                        :::||||                                   |  :||
Db        783 CWAEDPQERPPFQQIRLTLRKFNRENSSNILDNLLSRMEQYANNLEELVEERTQAYLEEK 842

Qy        388 KKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIV 447
              :| : ||| :|| ||| :|:    | |: :|:||| || |||| |  ::       |::|
Db        843 RKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAEST----PMQVV 898

Qy        448 NLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIH----HARSICHLALDM 503
              ||||||| || : |   | ||||||:|| || |||||   :     ||  :  :|| :
Db        899 TLLNDLYTCFDAVID---NFDVYKVETIGDAYMVVSGLP---VRNGRLHACEVARMALAL 952

Qy        504 MEIAGQVQV---DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKG 560
              ::     ::     | ::: |||||| |  ||:| :|||||||:|||  || |: ||
Db        953 LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGEAL 1012

Qy        561 KINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSRKNTGT 611
```

```
                ||::|  |   |         | || || | |||| :    | |  : : |
Db       1013 KIHLSSETKAVL----EEFGGFELELRGDVEMKGKGKVRTYWLLGERGSST 1059


RESULT 13
AAU11281
ID   AAU11281 standard; Protein; 1061 AA.
XX
AC   AAU11281;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human atrionatriuretic peptide receptor A (NPR1) variant protein.
XX
KW   Human; natriuretic peptide receptor A/guanylate cyclase A; NPR1;
KW   atrionatriuretic peptide receptor A; haplotyping; cytostatic; genotyping;
KW   haplotype pair; single nucleotide polymorphism; gene therapy;
KW   drug screening; hypertension; hypotensive.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "Pro substituted by Leu"
FT   Misc-difference 6
FT                   /note= "Arg substituted by Ser"
FT   Misc-difference 182
FT                   /note= "Ala substituted by Val"
FT   Misc-difference 341
FT                   /note= "Met substituted by Ile"
XX
PN   WO200179231-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US12300.
XX
PR   14-APR-2000; 2000US-197330P.
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Bentivegna SC,  Choi JY,  Kliem SE,  Nandabalan K;
XX
DR   WPI; 2002-066340/09.
XX
PT   Genotyping human natriuretic peptide receptor A/guanylate cyclase gene
PT   of an individual, involves determining identity of nucleotide pair at
PT   specific polymorphic sites for two copies of the gene  -
XX
PS   Disclosure; Page -; 96pp; English.
XX
CC   The invention relates to single nucleotide polymorphisms in the gene
CC   encoding the human natriuretic peptide receptor A/guanylate cyclase A
CC   (atrionatriuretic peptide receptor A) or NPR1 polypeptide. A method for
CC   haplotyping the NPR1 gene in an individual comprises identifying the
CC   nucleotide at one or more polymorphic sites and determining whether one
CC   of the copies of the gene is defined by one of the NPR1 haplotypes given
CC   in the specification or whether both copies are defined by a haplotype
CC   pair. This method is useful in genotyping, whereby all possible haplotype
CC   pairs can be assigned to specific genotypes. An association between a
CC   trait and a haplotype or haplotype pair of the NPR1 gene can be
CC   identified by comparing the frequency of the haplotype or haplotype pair
CC   in a population exhibiting the trait with the frequency of the haplotype
CC   or haplotype pair in a reference population, where a higher haplotype
CC   frequency in the trait population indicates the trait is associated with
CC   the haplotype or haplotype pair. NPR1 and its corresponding DNA are used
CC   for studying the expression and function of NPR1, for use in screening
CC   for candidate drugs to treat diseases related to NPR1 activity, such as
CC   hypertension. The sequences are also useful for studying the effect of
CC   variation on the biological activity of NPR1 as well as on the binding
CC   affinity of candidate drugs targeting NPR1. This sequence represents a
CC   human NPR1 variant polypeptide.
XX
```

```
SQ   Sequence   1061 AA;

  Query Match            14.9%;  Score 481.5;  DB 23;  Length 1061;
  Best Local Similarity   27.7%;  Pred. No. 3.6e-37;
  Matches  197;  Conservative   88;  Mismatches  227;  Indels  199;  Gaps   31;

Qy         33 EEGQFLVRIIYDDSKTYDLVAAASKVLN-------------------------LNAGEI 66
              | | | | : |: : : :||| : : ||                           |: |:
Db        416 ENGAFRVVLNYNGT-SQELVAVSGRKLNWPLGYPPPDIPKCGFDNEDPACNQDHLSTLEV 474

Qy         67 LQMFGKM---------FFVF----CQESGYDTILRVLGSNVREFLQNLDALHDHLATIYP 113
              | : | :         ||::     ::     : ||   :|        :|  ||
Db        475 LALVGSLSLLGILIVSFFIYRKMQLEKELASELWRVRWEDVEP-----SSLERHL----- 524

Qy        114 GMRAPSFRCTDAEKGK--GLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRN 171
                |:   | | : :|   | :|      ||   :     ||   :  | : :| : ::
Db        525 --RSAGSRLTLSGRGSNYGSLL----TTEGQFQV---FAKTAYYK--GNLVAVKRVNRKR 573

Qy        172 EECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYT-FCKAFPPHIIFDRDLVVT 230
              |   |: :: | :   :   | | ||    |    |   | :|       | : : :
Db        574 IEL--TRKVLFELKHMRDVQNEHLTRFVGACTDPPNICILTEYCPRGSLQDILENESITL 631

Qy        231 ------QCGNAIYRVLPQLQPGN-CSLLSVFSLVRPHIDISFHGILSHINTV----FVLR 279
                   | | : :  |  |  ||                 || |   | |    |||:
Db        632 DWMFRYSLTNDIVKGMLFLHNGAICS----------------HGNLKSSNCVVDGRFVLK 675

Qy        280 SKE-GLLDVEKLECEDELTGTEISCLRLKGQMIY---LPEADSILFLCSPSV--MNLDDL 333
               : ||     |: |              :|  :|   |  |  :| : || |      |:
Db        676 ITDYGLESFRDLDPE------------QGHTVYAKKLWTAPELLRMASPPVRGSQAGDV 722

Qy        334 TRRGLYLSDIPL-----HDATRDL--------VLLGEQ--FREEYKLTQELEIL------ 372
               |: | :| |      |    ||         |  |||  ||    |   || |
Db        723 YSFGIILQEIALRSGVFHVEGLDLSPKEIIERVTRGEQPPFRPSLALQSHLEELGLLMQR 782

Qy        373 -----------TDRLQLTLR---------------------------------ALEDEK 387
                        :::||||                                   |  :||
Db        783 CWAEDPQERPPFQQIRLTLRKFNRENSSNILDNLLSRMEQYANNLEELVEERTQAYLEEK 842

Qy        388 KKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIV 447
              :| : ||| :|| ||| :|:    | |: :|:||| || |||| |  ::       |::|
Db        843 RKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAEST----PMQVV 898

Qy        448 NLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIH----HARSICHLALDM 503
              ||||||| || : |    |  ||||||:|| || |||||   :     ||  :  :|| :
Db        899 TLLNDLYTCFDAVID---NFDVYKVETIGDAYMVVSGLP---VRNGRLHACEVARMALAL 952

Qy        504 MEIAGQVQV---DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKG 560
              ::     ::     | ::: |||||| |  ||:| :|||||||:|||  || |: ||
Db        953 LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGEAL 1012

Qy        561 KINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSRKNTGT 611
              ||::|  |   |         | || || | |||| :    | |  : : |
Db       1013 KIHLSSETKAVL----EEFGGFELELRGDVEMKGKGKVRTYWLLGERGSST 1059


RESULT 14
ABB11783
ID   ABB11783 standard; peptide; 1075 AA.
XX
AC   ABB11783;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human ANP-A receptor homologue, SEQ ID NO:2153.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03800.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   N-PSDB; ABA09027.
XX
PT   Human proteins and DNA encoding sequences useful for preventing,
PT   treating or ameliorating a medical condition in a mammalian subject
PT   e.g. arthritis and cancer -
XX
PS   Claim 20; Page 246-247; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell proliferation or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g., by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g., of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a novel human
CC   polypeptide of the invention.
XX
SQ   Sequence   1075 AA;

  Query Match            14.9%;  Score 481.5;  DB 22;  Length 1075;
  Best Local Similarity   27.7%;  Pred. No. 3.7e-37;
  Matches  197;  Conservative   88;  Mismatches  227;  Indels  199;  Gaps   31;
```

```
Qy          33 EEGQFLVRIIYDDSKTYDLVAAASKVLN-------------------------LNAGEI 66
Db         430 ENGAFRVVLNYNGT-SQELVAVSGRKLNWPLGYPPPDIPKCGFDNEDPACNQDHLSTLEV 488

Qy          67 LQMFGKM---------FFVF----CQESGYDTILRVLGSNVREFLQNLDALHDHLATIYP 113
Db         489 LALVGSLSLLLGILIVSFFIYRKMQLEKELASELWRVRWEDVEP-----SSLERHL----- 538

Qy         114 GMRAPSFRCTDAEKGK--GLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRN 171
Db         539 --RSAGSRLTLSGRGSNYGSLL----TTEGQFQV---FAKTAYYK--GNLVAVKRVNRKR 587

Qy         172 EECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVT 230
Db         588 IEL--TRKVLFELKHMRDVQNEHLTRFVGACTDPPNICILTEYCPRGSLQDILENESITL 645

Qy         231 ------QCGNAIYRVLPQLQPGN-CSLLSVFSLVRPHIDISFHGILSHINTV----FVLR 279
Db         646 DWMFRYSLTNDIVKGMLFLHNGAICS---------------HGNLKSSNCVVDGRFVLK 689

Qy         280 SKE-GLLDVEKLECEDELTGTEISCLRLKGQMIY---LPEADSILFLCSPSV--MNLDDL 333
Db         690 ITDYGLESFRDLDPE------------QGHTVYAKKLWTAPELLRMASPPVRGSQAGDV 736

Qy         334 TRRGLYLSDIPL-----HDATRDL--------VLLGEQ--FREEYKLTQELEIL------ 372
Db         737 YSFGIILQEIALRSGVFHVEGLDLSPKEIIERVTRGEQPPFRPSLALQSHLEELGLLMQR 796

Qy         373 -----------TDRLQLTLR---------------------------------ALEDEK 387
Db         797 CWAEDPQERPPFQQIRLTLRKFNRENSSNILDNLLSRMEQYANNLEELVEERTQAYLEEK 856

Qy         388 KKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIV 447
Db         857 RKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAEST----PMQVV 912

Qy         448 NLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIH----HARSICHLALDM 503
Db         913 TLLNDLYTCFDAVID---NFDVYKVETIGDAYMVVSGLP---VRNGRLHACEVARMALAL 966

Qy         504 MEIAGQVQV---DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKG 560
Db         967 LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGEAL 1026

Qy         561 KINVSEYTYRCLMSPENSDPQPFHLEHRGPVSMKGKKEPMQVWFLSRKNTGT 611
Db        1027 KIHLSSETKAVL----EEFGGFELELRGDVEMKGKGKVRTYWLLGERGSST 1073
```

RESULT 15
AAR38862
ID AAR38862 standard; Protein; 1029 AA.
XX
AC AAR38862;
XX
DT 08-FEB-1994 (first entry)
XX
DE GC-A.
XX
KW Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
KW mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
KW binding domain; antibody; diarrhoea; ligand; antagonist.
XX
OS Rattus rattus.
XX
PN US5237051-A.
XX
PD 17-AUG-1993.
XX
PF 06-DEC-1990; 90US-0623033.
XX
PR 06-DEC-1990; 90US-0623033.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Garbers DL, Schulz S;
XX
DR WPI; 1993-272183/34.
XX
PT New purified enterotoxin receptor protein - used to develop
PT prods. for treating abnormal conditions caused by bacterially
PT released enterotoxin, partic. diarrhoea
XX
PS Disclosure; Fig 3; 26pp; English.
XX
CC The sequences given in AAR38862-63 represent the guanylyl cyclases,
CC GC-A and GC-B, which binds heat stable enterotoxin. These proteins
CC are enterotoxin receptors which may be used as a therapeutic to control
CC intestinal fluid permeation as well as abnormal conditions caused
CC by bacterially released enterotoxin. The binding domain of the
CC proteins, or antibodies to the proteins, can be used to eliminate
CC diarrhoea. The proteins may be used to isolate ligands and to screen
CC for antagonists of toxin binding.
XX
SQ Sequence 1029 AA;

Query Match 14.5%; Score 468; DB 14; Length 1029;
Best Local Similarity 43.4%; Pred. No. 7.2e-36;
Matches 112; Conservative 38; Mismatches 82; Indels 26; Gaps 9;

```
Qy         361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
Db         789 EQYANNLE-ELVEERTQAYL----EEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSV 843

Qy         421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
Db         844 TIYFSDIVGFTALSAEST----PMQVVTLLNDLYTCFDAVID---NFDVYKVETIGDAYM 896

Qy         481 TVSGLPEPCIH----HARSICHLALDMMEIAGQVQV---DGESVQITIGIHTGEVVTGVI 533
Db         897 VVSGLP---VRNGQLHAREVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVV 953

Qy         534 GQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMK 593
Db         954 GLKMPRYCLFGDTVNTASRMESNGEALKIHLSSETKAVL---EEFD-GFELELRGDVEMK 1009

Qy         594 GKKEPMQVWFLSRKNTGT 611
Db        1010 GKGKVRTYWLLGERGCST 1027
```

Search completed: June 27, 2003, 13:03:03
Job time : 35.6287 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:07 ; Search time 17.4966 Seconds
(without alignments)
3401.080 Million cell updates/sec

Title: US-09-762-767A-4
Perfect score: 3231
Sequence: 1 MYGFVNHALELLVIRNYGPE.....QVWFLSRKNTGTEETKQDDD 619

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3231	100.0	619	1	S23097	guanylate cyclase
2	3206	99.2	619	1	OYBO70	guanylate cyclase
3	3188	98.7	619	1	OYRTB1	guanylate cyclase
4	987	30.5	682	1	OYRTB2	guanylate cyclase
5	960.5	29.7	685	2	T19968	hypothetical prote
6	955	29.6	686	2	T23721	hypothetical prote
7	938	29.0	751	2	T24458	hypothetical prote
8	914	28.3	699	2	T18984	hypothetical prote
9	909.5	28.1	732	2	S18325	guanylate cyclase,
10	793	24.5	690	1	OYRTA1	guanylate cyclase
11	770	23.8	691	1	OYBO77	guanylate cyclase
12	752.5	23.3	1099	2	T16822	hypothetical prote
13	730.5	22.6	717	2	S23098	guanylate cyclase
14	712	22.0	683	2	JH0810	guanylate cyclase
15	701.5	21.7	583	2	F88642	protein C54E4.3 [i
16	646.5	20.0	947	2	T20156	hypothetical prote
17	481.5	14.9	1061	1	OYHUAR	natriuretic peptid
18	470	14.5	1057	2	T15720	hypothetical prote
19	468	14.5	1057	1	OYRTR	atrial natriuretic
20	468	14.5	1057	2	I55319	guanylyl cyclase A
21	466	14.4	1057	2	I57963	natriuretic peptid
22	457	14.1	1057	1	OYMSAR	atrial natriuretic
23	456.5	14.1	632	2	T31667	guanylate cyclase
24	455.5	14.1	632	2	T31666	natriuretic peptid
25	454.5	14.1	1047	1	OYRTBR	atrial natriuretic
26	453.5	14.0	1047	1	OYHUBR	natriuretic peptid
27	451	14.0	1525	1	A56699	guanylate cyclase
28	450.5	13.9	1047	2	I45882	guanylate cyclase
29	449.5	13.9	1025	2	T21319	hypothetical prote
30	443	13.7	1056	2	T33167	hypothetical prote
31	438.5	13.6	1110	1	S55279	guanylate cyclase
32	438.5	13.6	1112	2	T28082	hypothetical prote
33	431.5	13.4	1005	2	S33525	guanylate cyclase
34	429	13.3	1012	2	T24384	hypothetical prote
35	428.5	13.3	1108	2	A55915	guanylate cyclase
36	425	13.2	1050	2	S45636	natriuretic-peptid
37	424	13.1	1068	2	T42382	guanylate cyclase
38	424	13.1	1108	2	I59385	guanylate cyclase
39	422	13.1	1102	2	JH0717	guanylate cyclase
40	420	13.0	1108	2	B55915	guanylate cyclase
41	419	13.0	1103	2	JC5581	guanylate cyclase
42	418	12.9	1122	2	T28130	hypothetical prote
43	416	12.9	1130	2	A89130	protein F52E1.4 [i
44	415.5	12.9	540	2	T34187	hypothetical prote
45	415.5	12.9	1110	2	I59370	guanylate cyclase

ALIGNMENTS

RESULT 1
S23097
guanylate cyclase (EC 4.6.1.2), soluble, 70K chain - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 20-Apr-2000
C;Accession: S23097
R;Giuili, G.; Scholl, U.; Bulle, F.; Guellaen, G.
FEBS Lett. 304, 83-88, 1992
A;Title: Molecular cloning of the cDNAs coding for the two subunits of soluble guanyl
A;Reference number: S23097; MUID:92316204; PMID:1352257
A;Accession: S23097
A;Molecule type: mRNA
A;Residues: 1-619 <GIU>
A;Cross-references: EMBL:X66533; NID:g31685; PIDN:CAA47144.1; PID:g31686
C;Genetics:
A;Gene: GDB:GUCY1B3; GUC1B3; GC-SB3
A;Cross-references: GDB:141992; OMIM:139397
A;Map position: 4q31.3-4q33
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: heterodimer; phosphorus-oxygen lyase
F;373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 100.0%; Score 3231; DB 1; Length 619;
Best Local Similarity 100.0%; Pred. No. 1.1e-215;
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              |||||||||||||||||||
Db        601 VWFLSRKNTGTEETKQDDD 619


RESULT 2
OYBO70
guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - bovine
N;Alternate names: guanylate cyclase, soluble, 70K chain
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
C;Accession: S01653; A38759
R;Koesling, D.; Herz, J.; Gausepohl, H.; Niroomand, F.; Hinsch, K.D.; Muelsch, A.; Boehm
FEBS Lett. 239, 29-34, 1988
A;Title: The primary structure of the 70 kDa subunit of bovine soluble guanylate cyclase
A;Reference number: S01653; MUID:89031214; PMID:2903071
A;Accession: S01653
A;Molecule type: mRNA
A;Residues: 1-619 <KOE>
A;Cross-references: EMBL:Y00770; NID:g407; PIDN:CAA68739.1; PID:g408
A;Accession: A38759
A;Molecule type: protein
A;Residues: 1-6;28-32;41-45;96-109;337-343;396-406;562-569 <KOE2>
C;Genetics:
A;Introns: 281/3; 392/2
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             99.2%;  Score 3206;  DB 1;  Length 619;
  Best Local Similarity   99.0%;  Pred. No. 5.8e-214;
  Matches  613;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCTDADKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKSECEDELTGTE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMTPENSDPQFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              |||||||||||||:||::
Db        601 VWFLSRKNTGTEETEQDEN 619


RESULT 3
OYRTB1
guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - rat
N;Alternate names: guanylate cyclase, soluble, 70K chain
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A31871
R;Nakane, M.; Saheki, S.; Kuno, T.; Ishii, K.; Murad, F.
Biochem. Biophys. Res. Commun. 157, 1139-1147, 1988
A;Title: Molecular cloning of a cDNA coding for 70 kilodalton subunit of soluble guan
A;Reference number: A31871; MUID:89087429; PMID:2905128
A;Accession: A31871
A;Molecule type: mRNA
A;Residues: 1-619 <NAK>
A;Cross-references: GB:M22562; NID:g204273; PIDN:AAA41204.1; PID:g204274
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             98.7%;  Score 3188;  DB 1;  Length 619;
  Best Local Similarity   98.5%;  Pred. No. 1e-212;
  Matches  610;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||||||||||||||||||||:|||||||||| ||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDQEGQFLVRIIYDHSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRSEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQDSRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||| | ||||
Db        241 PQLQPGKCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGAE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
```

```
Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              ||||||||||||| ||::
Db        601 VWFLSRKNTGTEETNQDEN 619


RESULT 4
OYRTB2
guanylate cyclase (EC~4.6.1.2), soluble, beta-2 chain - rat
N;Alternate names: guanylate cyclase, soluble, 76K chain
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A36228
R;Yuen, P.S.T.; Potter, L.R.; Garbers, D.L.
Biochemistry 29, 10872-10878, 1990
A;Title: A new form of guanylyl cyclase is preferentially expressed in rat kidney.
A;Reference number: A36228; MUID:91105012; PMID:1980215
A;Accession: A36228
A;Molecule type: mRNA
A;Residues: 1-682 <YUE>
A;Cross-references: GB:M57507; GB:J05308; NID:g204279; PIDN:AAA41207.1; PID:g204280
A;Experimental source: kidney
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;360-584/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             30.5%;  Score 987;  DB 1;  Length 682;
  Best Local Similarity   36.9%;  Pred. No. 2.4e-60;
  Matches  225;  Conservative  104;  Mismatches  195;  Indels   86;  Gaps   12;

Qy         66 ILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDA 125
              ||::||: || ||: |||| :|| || |: ||::|||||| :||  |  | ||||  :
Db          4 ILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDALHSYLALSYQEMNAPSFRVEEG 63

Qy        126 EKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC------DHTQF 179
                |  ::|||||:| ||  || |||: ||:    |:: | ::   |||        :|  |
Db         64 ADG-AMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMSILDM-NEEVERTGKKEHVVF 121

Qy        180 LIEEKESKE-----------EDFYED-------------------------------- 194
              |: :|  ::           ||   |
Db        122 LVVQKAHRQIRGAKASRPQGSEDSQADQEALQGTLLRMKERYLNIPVCPGEKSHSTAVRA 181

Qy        195 ---------LDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQP 245
                      | |:    :   :    || ||||||:||  | | | |  | : :| :
Db        182 SVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFHIVFDEALRVKQAGVNIQKYVPGILT 241

Qy        246 GNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLR 305
              :|   ||:: | :  :   |   ||: |||:::: ::   :           |:
Db        242 QKFALDEYFSIIHPQVTFNISSICKFINSQFVLKTRKEMMPKAR---------KSQPMLK 292

Qy        306 LKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKL 365
              |:|||::     ::|:|||:| :| :|    ::|||  || ||||:|| :|   | :|
Db        293 LRGQMIWMESLRCMIFMCSPNVRSLQELEESKMHLSDIAPHDTTRDLILLNQQRLAEMEL 352

Qy        366 TQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFS 425
              : :||   : |::    |  |||||:||||::|| |||:|:  | | |   ::  ||||
Db        353 SCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVANQLKEGRKVAAGEFETCTILFS 412

Qy        426 GIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGL 485
              :| |   |   |: |  ::|||:|| :|::|| ||       ||||||:|| || | |:
Db        413 DVVTFTNIC---AACE-PIQIVNMLNSMYSKFDRLTSVHD---VYKVETIGDAYMVVGGV 465

Qy        486 PEPCIHHARSICHLALDMMEIAGQVQ--VDGESVQITIGIHTGEVVTGVIGQRMPRYCLF 543
              | |   ||: : : || |   | :|   | || :|| :||||| |: ||:| :|||||||
Db        466 PVPVESHAQRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVVGDKMPRYCLF 525
```

```
Qy        544 GNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWF 603
              |:|||  || |: |   |:::|   :| |      :  | :  || : :||| : |  :|
Db        526 GDTVNTASRMESHGLPSKVHLSPTAHRAL-----KNKGFEIVRRGEIEVKGKGK-MTTYF 579

Qy        604 LSRKNTGTEE 613
              | :    ||:
Db        580 LIQNLNATED 589


RESULT 5
T19968
hypothetical protein C46E1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T19968; T23092
R;Lennard, N.
submitted to the EMBL Data Library, November 1997
A;Reference number: Z19205
A;Accession: T19968
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-685 <WIL>
A;Cross-references: EMBL:AL008867; PIDN:CAA15512.1; GSPDB:GN00028; CESP:C46E1.2
A;Experimental source: clone C46E1
R;Lennard, N.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z19673
A;Accession: T23092
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-685 <WI2>
A;Cross-references: EMBL:Z99942; PIDN:CAB17073.1; GSPDB:GN00028; CESP:C46E1.2
A;Experimental source: clone H13N06
C;Genetics:
A;Gene: CESP:C46E1.2
A;Map position: X
A;Introns: 11/3; 36/2; 70/1; 133/2; 179/1; 202/3; 258/2; 298/2; 301/3; 325/2; 405/2;
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match             29.7%;  Score 960.5;  DB 2;  Length 685;
  Best Local Similarity   34.2%;  Pred. No. 1.6e-58;
  Matches  220;  Conservative  131;  Mismatches  234;  Indels   59;  Gaps   17;

Qy          2 YGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNL 61
              :||:: ::  |:|| ||   |  : : |  :   : ::   | |: |: || | | :| :
Db         12 FGFIHESIRQLMIRTYGEAFWSKVLERAGFEAGKENIINHYYSDADTFSLVDAVSVILKV 71

Qy         62 NAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALH---DHLATIY-PGMRA 117
                 :: :|:|     :  |:|:| ::| :  |:: || |||:||   ||:  :|   :|
Db         72 TREQVWEMYGCFLIQYTMETGWDDLIRSMSPNLKGFLDNLDSLHYFIDHV--VYKANLRG 129

Qy        118 PSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVI--QQRNEEC- 174
              ||||| |   |  : ||||: | ||  || |:::  |:::   :: | :    ||: :
Db        130 PSFRCEDNPDGT-ITLHYYTGRPGLYPIVKGVLREAAKRVFKLDVSMTITGRTQRSVQMA 188

Qy        175 ------DHTQFLIE---EKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDR 225
                   :|  ||::     :| ||     :  |   : |::   |   ||:|:: |:
Db        189 TGERIEEHVIFLVKTLNTDQSNEEALGTAVVQHSNN--YKIRLTHMDFISTFPYHMVVDQ 246

Qy        226 DLVVTQCGNAIYRVLPQ-LQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGL 284
              |  : | |  :|  :|: |       |: :| : || | : |  | : || ||||: |
Db        247 DCKIVQVGRELYNHIPKDLLSVGTPLMRIFEVTRPQIPLDFDSICNFINAVFVLQVKTTP 306

Qy        285 LDVEK---------LECEDE--------LTGTEISCLRLKGQMIYLPEADSILFLCSPSV 327
              :: ::          :|  :          |  ::   |:|||||: :     |::|||| |
Db        307 MEFQRNANKRAAQAIEASENLYEDNNGALALSQSQHLKLKGQMMLMSSGGHIMYLCSPYV 366

Qy        328 MNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEK 387
              :: :| : || |: :|:|| ||||:|| :|   : ::  :||   ::|:   : || ||
Db        367 TSIPELLQYGLRLTAMPIHDPTRDLILLNQQRLSDVEMNLQLEANNEQLENMAKDLEVEK 426

Qy        388 KKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFN---AFCSKHASGEGAM 444
```

```
Db        427 GKTDALLREMLPPSVAQQLKQGLSVEAREYEEATVMFTDVPTFQQIVPLCT-------PK 479

Qy        445 KIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMM 504

Db        480 DIVHLLNELFTKFDRLIGIQK---AYKVETVGDSYMSVGGIPDLVDDHCEVICHLALGMV 536

Qy        505 EIAGQV--QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKI 562

Db        537 MEARTVCDPITNTPLHIRAGIHSGPVVAGVVGAKMPRYCLFGDTVNTSSRMESHSPIGRI 596

Qy        563 NVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSR 606

Db        597 HCSENAKKC----AESTGRFEFEPRGRVQIKGKGE-MNTYFLLR 635


RESULT 6
T23721
hypothetical protein M04G12.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T23721
R;Sims, M.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19789
A;Accession: T23721
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-686 <WIL>
A;Cross-references: EMBL:Z81103; PIDN:CAB03210.1; GSPDB:GN00023; CESP:M04G12.3
A;Experimental source: clone M04G12
C;Genetics:
A;Gene: CESP:M04G12.3
A;Map position: 5
A;Introns: 1/3; 26/2; 60/1; 192/3; 255/2; 321/2; 401/2; 645/1
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match             29.6%;  Score 955;  DB 2;  Length 686;
  Best Local Similarity   34.2%;  Pred. No. 4e-58;
  Matches  230;  Conservative  126;  Mismatches  235;  Indels   82;  Gaps   20;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Db          1 MFGFIHESIRQLVIRKYGEDVWLQVLERSGFENGKENIVNHYYSDTDTYVLVDSVSIVLK 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALH---DHLATIY-PGMR 116

Db         61 VTKDQIWEMYGGFLITYSMEIGWDELVRSMSPNLKGFLDNLDSLHYFIDHV--VYKANLR 118

Qy        117 APSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC-- 174

Db        119 GPSFRCEENPDGT-LMLHYFTGRPGLYHIVKGVVKEVAKLVFNLDISL-VVQGRTQRSVH 176

Qy        175 --------DHTQFLIEEKE---------------SKEEDFYEDLDRFEENGTQESRISPY 211

Db        177 MNNGERVEEHVIFLIKNVEEPRRDSDTSTTSALTSVEPDFGEIID-------DNLKVSLQ 229

Qy        212 TFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQ--LQPGNCSLLSVFSLVRPHIDISFHGIL 269

Db        230 DFSRALPYHFVLDESCRLVQCGDELYNHIPNELLQPGT-PILRIFEINRPQIPLDFENIC 288

Qy        270 SHINTVFVLRSKEGLL---------------DVEKLECEDELTGTEISC-LRLKGQMIYL 313

Db        289 NFINAVFVLQVKTSPLRKKHMNAMTKEEREQEVEAMEEEVESNELTQGCHLKLKGQMMML 348

Qy        314 PEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILT 373

Db        349 STKKHIIYLCSPYVTSINELMQFGMRLTAMPLHDATRDLILLNQQRLSDVEVNLQLEANN 408

Qy        374 DRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAF 433

Db        409 EQLETMTHELEVERQKTDSILKDMLPRKIAKQLLSGEHLEPCEYE-ATVMFCDLPAFQQI 467

Qy        434 ---CSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490

Db        468 IPVCQ-------PKNIVKLLNEVFFKLDRIVVLRG---VYKVETVSDSYMTVSGIPDYTS 517

Qy        491 HHARSICHLALDMMEIAGQVQ--VDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVN 548

Db        518 EHAENMCHVALGMMWEARSVMDPVNKTPFLLRIGLHSGTIIAGVVGTKMPRYCLFGETVT 577

Qy        549 LTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSR-- 606

Db        578 LASQMESLGVAGKIQCSSWTY----SKAMETGRFEFSPRGRINVKGRGD-VETYFLMRSL 632

Qy        607 KNTGTEETKQDDD 619

Db        633 KKSIWEITDHERD 645


RESULT 7
T24458
hypothetical protein T04D3.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T24458
R;Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19893
A;Accession: T24458
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-751 <WIL>
A;Cross-references: EMBL:Z81114; PIDN:CAB03288.1; GSPDB:GN00019; CESP:T04D3.4
A;Experimental source: clone T04D3
C;Genetics:
A;Gene: CESP:T04D3.4
A;Map position: 1
A;Introns: 1/3; 26/2; 95/3; 149/1; 213/3; 247/3; 291/3; 332/3; 369/2; 418/1; 462/3; 4
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match             29.0%;  Score 938;  DB 2;  Length 751;
  Best Local Similarity   35.0%;  Pred. No. 6.8e-57;
  Matches  220;  Conservative  125;  Mismatches  236;  Indels   48;  Gaps   15;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Db          1 MFGWIHESFRQLVTRKYGKDIWEKIVHMSKFELGTESEIAHYYNDDETLRLVNSMANVIG 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALH---DHLATIY-PGMR 116

Db         61 IPIEEIWEAYGGFLIQFTMETGWDELLRAMAPDLEGFLDSLDSLHYFIDHV--VYKTKLR 118

Qy        117 APSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECD- 175

Db        119 GPSFRC-DVQADGTLLLHYYSKRSGLYPIVKGVVREVARRIYDTEVVMKVQERKQEHLDA 177

Qy        176 ----HTQFLIEEKE----------SKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHI 221

Db        178 FVTEHVVFVITQIENANSTQPKSISSKADSQIDL----STGIYE--ISSSDFSLAFPYHI 231

Qy        222 IFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSK 281

Db        232 CFDPDLFVEHFGNFIKKTFPNAMRQETRVTDLLELVHPEVPFSYESIKYYKNSLFVFRLK 291

Qy        282 EGLLDVEKLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLS 341

Db        292 -GLGDIV------HNANDEAKTVLLKGSMVFIDEGKYILYMCSVNVTTVRELIERNLHLS 344

Qy        342 DIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPS 401

Db        345 DMQRHDGTRDVIMLNQSRMSQVELNRTLEETTKKLKKMAQELEIEKQKTDELLCELMPAS 404

Qy        402 VANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLT 461

Db        405 VADSLRSGKAMDAKEFADCTLLFTDIVTFTNICAMCTPYD----VVTLLNDLYLRFDRLV 460
```

```
Qy        462 DSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV--QVDGESVQI 519
                  |||||:|| || | |:|| | :|| : :::: |:  :  |   :  : ::|
Db        461 GLHD---AYKVETIGDAYMIVGGVPERCENHAERVLNISIGMLMESKLVLSPITHKPIKI 517

Qy        520 TIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSD 579
              :|:| | || ||:| :||||||||:|||: :: |: | : ||:|||      ::   ::
Db        518 RLGVHCGPVVAGVVGIKMPRYCLFGDTVNVANKMESNGIQCKIHVSE---TGKLNGLKAN 574

Qy        580 PQFHLEHRGPVSMKGKKEPMQVWFLSRKN 608
              | :    ||   ::| |  |  :|| | :
Db        575 PSYVFIDRGNTEIRG-KGMMYTYFLERND 602


RESULT 8
T18984
hypothetical protein C06B3.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T18984
R;Percy, C.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19056
A;Accession: T18984
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-699 <WIL>
A;Cross-references: EMBL:Z77652; PIDN:CAB01118.1; GSPDB:GN00023; CESP:C06B3.8
A;Experimental source: clone C06B3
C;Genetics:
A;Gene: CESP:C06B3.8
A;Map position: 5
A;Introns: 1/3; 60/1; 192/3; 210/2; 270/2; 336/2; 415/2; 659/1
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

  Query Match             28.3%;  Score 914;  DB 2;  Length 699;
  Best Local Similarity   32.9%;  Pred. No. 2.8e-55;
  Matches  219;  Conservative  132;  Mismatches  234;  Indels   80;  Gaps   19;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              |:||:: ::  ||||||| : |  : : :  :   : ::   | |: || || : | ||
Db          1 MFGFIHESIRQLVIRNYGEDTWTQVLERSGFESGKENIMNHYYSDTDTYVLVDSVSLVLK 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALH---DHLATIY-PGMR 116
              :   :: :|:|     :  | |:| ::| :  |:: || |||:||   ||:  :|   :|
Db         61 VTKDQVWEMYGGFLITYSMEIGWDELVRSMSPNLKGFLDNLDSLHYFIDHV--VYKANLR 118

Qy        117 APSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC-- 174
               ||||| :   |  |:|||:: | ||  || |::| ||:::   :| : |:| | :
Db        119 GPSFRCEETPDGT-LLLHYFTGRPGLYHIVKGVVKEVAKRVFDLDITL-VVQGRTQRSVH 176

Qy        175 --------DHTQFLIEEKESKEEDFYEDL------DRFEENGTQES-------------- 206
                     :|  ||| :|  :   |   :        | :  |:: |
Db        177 MNNGERVEEHVVFLINKKGGQFHPFLIAIVVTGSEPRRDSEGSEVSLLTSTNANFPTIVD 236

Qy        207 ---RISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQ--LQPGNCSLLSVFSLVRPHI 261
                 ||   | || |:| : |    : |||: ::  :|   ||||   :| :| : || |
Db        237 DTLGISLDDFSKALPYHFVIDESCKLVQCGSELHNHIPNELLQPGT-PILRIFEINRPQI 295

Qy        262 DISFHGILSHINTVFVLRSKEGLLD------------VEKLECEDELTGTEIS---CLRL 306
               : |  | : || |||| |   |               :::|  ||    |::    |:|
Db        296 PLDFENICNFINAVFVLQVKTSPLKKKHMDAMSQEELKQEMETLDEDATNELTQGHHLKL 355

Qy        307 KGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLT 366
              ||||: |      |::|||| | ::::| : |: |: :|||||||||:|| :|   : ::
Db        356 KGQMMLLASKKHIIYLCSPYVTSINELMQYGMRLTAMPLHDATRDLILLNQQRLSDVEVN 415

Qy        367 QELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSG 426
               :||   ::|:   | || |::|||::|  :||  :| :|    : |  ::  |::|
Db        416 LQLEANNEQLETMTRELELERQKTDSILKDMLPRRIAQQLLSGEHIEACEHE-ATVMFCD 474

Qy        427 IVGFNAF---CSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVS 483
              :  |      ||          |||:||::: : | :  |    |||||| | || ||
Db        475 LPAFQQAIPQCS-------PKDIVNMLNEIFRKLDRIVVIRG---VYKVETVSDSYMAVS 524

Qy        484 GLPEPCIHHARSICHLALDMMEIAGQV--QVDGESVQITIGIHTGEVVTGVIGQRMPRYC 541
              |:|:    || ::||:|| ||  |  |    |       : ||||:| :  ||:|   |:||
Db        525 GIPDYTPEHAENMCHVALGMMWEARSVIDPVSKTPFLLRIGIHSGTITAGVVGTVHPKYC 584

Qy        542 LFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQV 601
              ||| || | |: |: |  |||  |:: |:  |       :|    || : :| ::   :
Db        585 LFGETVTLASQMESLGMAGKIQCSKWAYQKAM----ETGRFEFSPRGRIDVK-QRGLTET 639

Qy        602 WFLSR 606
              :||:|
Db        640 YFLTR 644


RESULT 9
S18325
guanylate cyclase, soluble, alpha chain - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S18325
R;Harteneck, C.; Wedel, B.; Koesling, D.; Malkewitz, J.; Boehme, E.; Schultz, G.
FEBS Lett. 292, 217-222, 1991
A;Title: Molecular cloning and expression of a new alpha-subunit of soluble guanylyl
A;Reference number: S18325; MUID:92070494; PMID:1683630
A;Accession: S18325
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-732 <HAR>
A;Cross-references: EMBL:X63282; NID:g31670; PIDN:CAA44921.1; PID:g31671
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
F;473-701/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             28.1%;  Score 909.5;  DB 2;  Length 732;
  Best Local Similarity   37.7%;  Pred. No. 6.1e-55;
  Matches  208;  Conservative  111;  Mismatches  175;  Indels   57;  Gaps   14;

Qy         51 LVAAASKVLNLNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLAT 110
               ::   : :| |   || : ||: ||  |     : :|| :|  :::|    ||| :|: |
Db        154 ILQCTANILGLKFEEIQKRFGEEFFNICFHEN-ERVLRAVGGTLQDFFNGFDALLEHIRT 212

Qy        111 IY---PGMRAPSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVI 167
               :     : :||| | :  :|  |:|||:     :   ::|:||   ::|:   :|::|
Db        213 SFGKQATLESPSFLCKELPEGT-LMLHYFHPHHIVGFAMLGMIKAAGKKIY--RLDVEVE 269

Qy        168 QQRNEE----------CDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAF 217
              |  ||:           |   |||  || :  :  ::| :        : |||  |||:||
Db        270 QVANEKLCSDVSNPGNCSCLTFLI--KECENTNIMKNLPQGTSQVPADLRISINTFCRAF 327

Qy        218 PFHIIFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSV------FSLVRPHIDISFHGILSH 271
              |||::||  : | | |  : : |       |    |       | :| | :: :|  :|
Db        328 PFHLMFDPSMSVLQLGEGLRKQL------RCDTHKVLKFEDCFEIVSPKVNATFERVLLR 381

Qy        272 INTVFVLRSKEGLLDVEKLECEDELTGTE--ISCLRLKGQMIYLPEADSILFLCSPSVMN 329
              ::| ||:|:|             | :|:|     : :|||||::|||::|||| || |
Db        382 LSTPFVIRTK-----------PEASGSENKDKVMEVKGQMIHVPESNSILFLGSPCVDK 429

Qy        330 LDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKK 389
              ||:|  |||:|||||:||||||::|:||| : :  | : :: |   |: | :|||:||||
Db        430 LDELMGRGLHLSDIPIHDATRDVILVGEQAKAQDGLKKRMDKLKATLERTHQALEEEKKK 489

Qy        390 TDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNL 449
              |  ||||: |   || :|   : | |:::|:||:||| |||| | |::       |:::::
Db        490 TVDLLYSIFPGDVAQQLWQGQQVQARKFDDVTMLFSDIVGFTAICAQCT----PMQVISM 545

Qy        450 LNDLYTRFDTLTDSRKNPF--VYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIA 507
              ||:|||||       :  |  :|||||:|| |   :||    : ||: |  :|| |||::
Db        546 LNELYTRFD-----HQCGFLDIYKVETIGDAYCVAAGLHRKSLCHAKPIALMALKMMELS 600

Qy        508 GQVQV-DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSE 566
               :|   ||  :|: ||||:| |: ||:| ||||||||| | | |: |:     :||||
Db        601 EEVLTPDGRPIQMRIGIHSGSVLAGVVGVRMPRYCLFGNNVTLASKFESGSHPRRINVSP 660
```

```
Qy        567 YTYRCLMSPEN 577
                ||: |   |:
Db        661 TTYQLLKREES 671
```

RESULT 10
OYRTA1
guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - rat
N;Alternate names: guanylate cyclase, soluble, 77K chain
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A38297
R;Nakane, M.; Arai, K.; Saheki, S.; Kuno, T.; Buechler, W.; Murad, F.
J. Biol. Chem. 265, 16841-16845, 1990
A;Title: Molecular cloning and expression of cDNAs coding for soluble guanylate cyclase
A;Reference number: A38297; MUID:91009100; PMID:1698769
A;Accession: A38297
A;Molecule type: mRNA
A;Residues: 1-690 <NAK>
A;Cross-references: GB:M57405; GB:M36075; NID:g204277; PIDN:AAA41206.1; PID:g204278
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;432-660/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match 24.5%; Score 793; DB 1; Length 690;
  Best Local Similarity 34.7%; Pred. No. 6.5e-47;
  Matches 214; Conservative 100; Mismatches 215; Indels 88; Gaps 20;

```
Qy         27 KEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEILQ-MFGKMFFVFCQESGYDT 85
              ||  :   | : | ||| :::     :||   |      |:|:   |:  |  | |   :
Db        102 KENRNSSEKEDLERIIAEEA-----IAAGVPV------EVLKDSLGEELFKICYEED-EH 149

Qy         86 ILRVLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSERE 140
              || |:|  :::|| :   |     |   |     :   |  | | ::   | ::|:  :
Db        150 ILGVVGGTLKDFLNSFSTLLKQSSHCQEAERRGRLEDASILCLDKDQ-DFLNVYYFFPKR 208

Qy        141 GLQDIVIGIIKTVAQQIHGTEIDMKVIQQ-RNEECDHTQFL--------IEEKESKEEDF 191
                 :: ||||  |: :: : ::: ::     ||  |:|:        :  | :|
Db        209 TTALLLPGIIKAAARILYESHVEVSLMPPCFRSEC--TEFVNQPYLLYSVHVKSTKPS-- 264

Qy        192 YEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCS 249
                  | : | : ||| |||| : | || | |:: :   |  :
Db        265 ------LSPGKPQSSLVIPTSLFCKTFPFHFMLDRDLAILQLGNGIRRLVNKRDFQGKPN 318

Qy        250 LLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQ 309
               | :: | |: :| ||:: :|  ||:| :             | |       : ||||
Db        319 FEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRW----------DNLVKKSSRVMDLKGQ 368

Qy        310 MIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQEL 369
              |||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        369 MIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRL 428

Qy        370 EILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVG 429
               |   |:   :|||:|||||  || |: |  || :|   : | ||::: ||:||| |||
Db        429 GKLKATLEHAHQALEEEKKKTVDLLCSIFPSEVAQQLWQGQIVQAKKFNEVTMLFSDIVG 488

Qy        430 FNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPC 489
              | | ||: :     :::: :|| ||||||          ||||||:|| |    ||
Db        489 FTAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHRES 541

Qy        490 IHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVN 548
               ||  |  :|| |||:: :|    || ::: ||:|:| |  ||:| :||||||||| |
Db        542 DTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVT 601

Qy        549 LTSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQF--------HL----EH 586
              | :: |:     ||||| ||| |          | |   | |        |     :|
Db        602 LANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLDAYQH 661

Qy        587 RGPVSMKGKKEPMQVWF 603
              :|| |        : ||
Db        662 QGPNS--------KPWF 670

RESULT 11
OYBO77
guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - bovine
N;Alternate names: guanylate cyclase, soluble, 77K chain
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
C;Accession: S10713; A38767
R;Koesling, D.; Harteneck, C.; Humbert, P.; Bosserhoff, A.; Frank, R.; Schultz, G.; B
FEBS Lett. 266, 128-132, 1990
A;Title: The primary structure of the larger subunit of soluble guanylyl cyclase from
A;Reference number: S10713; MUID:90306336; PMID:1973124
A;Accession: S10713
A;Molecule type: mRNA
A;Residues: 1-691 <KOE1>
A;Cross-references: EMBL:X54014; NID:g405; PIDN:CAA37960.1; PID:g406
A;Accession: A38767
A;Molecule type: protein
A;Residues: 118-133;226-232;286-293;319-330;412-417;557-571;629-637 <KOE2>
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase
F;434-662/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match 23.8%; Score 770; DB 1; Length 691;
  Best Local Similarity 33.9%; Pred. No. 2.5e-45;
  Matches 205; Conservative 90; Mismatches 200; Indels 110; Gaps 16;

```
Qy         27 KEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDT 85
              ||  :   | :   :|: | :     :||   |      ||: :  |:  |  | |   :
Db        104 KENRKSLEREDFEKIVVDQA-----IAAGVPV------EIIKESLGEELFKICYEED-EY 151

Qy         86 ILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDAEKGKG--------------- 130
              || |:|  :::|| :   |                |  | :||| ||
Db        152 ILGVVGGTLKDFLNSFSTL----------LKQSSHCQEAEK-KGRFEDASILCLDKDPD 199

Qy        131 -LILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFLIEEKESKEE 189
               | ::|:  :     |: ||||  |: :: ||:::     |
Db        200 VLYVYYFFPKRITSLILPGIIKAAARILYETEVEVSSTPSR---------------- 240

Qy        190 DFYEDLDRFEENGTQ----------------ESRISPYT-FCKAFPFHIIFDRDLVV 229
               |::|   | :   :                | : | : ||| |||| : |||: :
Db        241 -FHQDCREFVDQPCELYSVHIRSARPHPPPGKPVSSLVIPASLFCKTFPFHFMLDRDMSI 299

Qy        230 TQCGNAIYRVLPQLQ-PGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVE 288
               | |: | |:: :    |       | :: | |  :| ||:: :|  |::| :
Db        300 LQLGHGIRRLMSRRDVQGKPHFDEYFEILTPKISQTFSGIMTMLNMQFLVRVRRW----- 354

Qy        289 KLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDA 348
                 |          : |||||||: |: ||||| || |  |:| | ||||||||:|:|
Db        355 -----DNSMKKSSRVMDLKGQMIYMVESSSILFLGSPCVDRLEDFTGRGLYLSDIPIHNA 409

Qy        349 TRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRH 408
              ||:||:||| | :  | : |  |    |:   :|||:||:||  || |: |  || :|
Db        410 LRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKRKTVDLLCSIFPSEVARQLWQ 469

Qy        409 KRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPF 468
               | |||: |||:||| |||| | ||: :     :::: :|| ||||||
Db        470 GHAVQAKRFGNVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFDRQCGELD--- 522

Qy        469 VYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGE 527
              ||||||:|| |    || :    ||  |  :|| |||:: : |   || ::: ||:|:|
Db        523 VYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSHEVVSPHGEPIKMRIGLHSGS 582

Qy        528 VVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL----------MSPEN 577
              |  ||:| :|||||||| | | :: |:     |||||  ||| |            | |
Db        583 VFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREE 642

Qy        578 SDPQF 582
                | |
Db        643 LPPNF 647
```

RESULT 12
T16822
```

```
hypothetical protein T07D1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16822
R;Leimbac, D.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid T07D1.
A;Reference number: Z18584
A;Accession: T16822
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1099 <LEI>
A;Cross-references: EMBL:U41531; NID:g1109802; PID:g1109803; PIDN:AAA83158.1; CESP:T07D1
C;Genetics:
A;Gene: CESP:T07D1.1
A;Introns: 26/3; 94/3; 144/2; 200/3; 239/2; 278/3; 344/1; 382/2; 425/3; 474/3; 523/3; 61

  Query Match             23.3%;  Score 752.5;  DB 2;  Length 1099;
  Best Local Similarity   29.9%;  Pred. No. 8e-44;
  Matches  211;  Conservative  124;  Mismatches  254;  Indels  117;  Gaps   24;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              |||:   :   :   ||   | ::|  : : ::  | :   : :  :: |: |   |
Db          1 MYGLIIDHIATYIKEKYGESTWSEVKFVSGVTDD-TFQMDKKFSEGLSHKLIWACHDVTG 59

Qy         61 LNAGEILQMFGKMFFVFCQESGYDT----------ILRVLGSNVREFLQNLDALHDHLAT 110
                |::   |  |: |  :  ::           :|||||    :||  || ||::|
Db         60 DPVDELMTNIGTSFYKFLTKFEFNKVYYVLFYIPYVLRVLGRTFPQFLNGLDNLHEYLRF 119

Qy        111 IYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQ-Q 169
              :| :: ||| | : |   || ||| |:| |     | | |: ::|::  ||: ::::  :
Db        120 TFPKLKPPSFYC-EHESRTGLTLHYRSKRRGFLHYVQGQIRNISQELFQTEVVIELLDIE 178

Qy        170 RNEECDHTQFLIEEKESKEEDFYEDLDRFEENGT----------QESRISPYTFCKAFPF 219
              :   :|    :          : :|| |   ||          :: :|:   |   |||
Db        179 HDLNLEHVIMRLH---------FNNLD-FNRQGTAYRNLNDSILEKVKITSDIFFDIFPF 228

Qy        220 HIIFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPHI------------DISF-- 265
              |:|:| : :   |  : ||:  :      :   | |:|| |              : ||
Db        229 IIVFNRGMRIRNIGIGLLRVMAGIV--GKKINQTFLLMRPFIRFRWEEVRGTNRNYSFDI 286

Qy        266 ---------------------HGILSHINTVFVLRS-------KEGLLDVEKLECEDELT 297
                                  |: | | :| | |       ::|:| | |    |::|
Db        287 ACDCFRNFFASPKKRVLKSVKSQIMLHSNNIFELISSDPIQEDEDGIL-VYKTTDVDQMT 345

Qy        298 --------GTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDAT 349
                     |     | ||||| |: | :|| |:  | : :|  : : ||:::|  |||::
Db        346 EERHRMGDGEREKFLSLKGQMFYMEEWESICFVGIPVMSHLPQMYKSGLFINDFALHDSS 405

Qy        350 RDLVLLGEQFREEYKLTQELEILTDR-LQLTLRALEDEKKKTDTLLYSVLPPSVANELRH 408
              |||||    |   | ||    |    | ::  :  |: |:::|| ||| :|| ||||:|||
Db        406 RDLVLASTQQSAELKLLLHQEAQKSRNMRENMNRLKKERRRTDKLLYQMLPKSVANQLRH 465

Qy        409 KRPVPA--KRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKN 466
                  |  :|:|:|||||: || |   ||        ::::  |  :|| || : |:
Db        466 GESAVACCERFDSVTILFTDIVEFTKMCSSLT----PLEVIEFLKVIYTNFDKIIDTHG- 520

Qy        467 PFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGE--SVQITIGI 523
              ||||||:|| || ||| |    | |  |   |   :  ||: | :: :   : |  |:
Db        521 --VYKVETIGDAYMVVSGAPTKTEHDAEFILDCASQFLVEAGKMVNMNNKIHKIDIRAGV 578

Qy        524 HTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDP-QF 582
              |:| || ||:|  |||||||| || : :: |      || ||| |:  :   | |||  :
Db        579 HSGSVVAGVVGLSMPRYCLFGETVYVANKMEQNSSPMKILVSETTHNKI---EESDPGLY 635

Qy        583 HLEHRGPVSMKGKKEPMQVWFL----------SRKNTGTEETKQDD 618
              | |  : :|   : :| :|:           | :|    |::|||
Db        636 QFERREEIEIK-DDQTIQTFFVVSRHGPHRVPSPRNC---ESRQDD 677


RESULT 13
S23098
guanylate cyclase (EC 4.6.1.2), soluble, 81K chain - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S23098
R;Giuili, G.; Scholl, U.; Bulle, F.; Guellaen, G.
FEBS Lett. 304, 83-88, 1992
A;Title: Molecular cloning of the cDNAs coding for the two subunits of soluble guanyl
A;Reference number: S23097; MUID:92316204; PMID:1352257
A;Accession: S23098
A;Molecule type: mRNA
A;Residues: 1-717 <GIU>
A;Cross-references: EMBL:X66534; NID:g31683; PIDN:CAA47145.1; PID:g31684
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: heterodimer; phosphorus-oxygen lyase
F;430-658/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             22.6%;  Score 730.5;  DB 2;  Length 717;
  Best Local Similarity   36.9%;  Pred. No. 1.5e-42;
  Matches  182;  Conservative   80;  Mismatches  178;  Indels   53;  Gaps   10;

Qy        119 SFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ 178
              |  | | |    | ::|:  :     |: ||||  |  :: ||::: ::      | |
Db        187 SILCLDKE-DDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLM----PPCFHND 241

Qy        179 ----------FLIEEKESKEEDFYEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRD 226
                       :  | :|             :  | | : ||| |||| :||:|
Db        242 CSEFVNQPYLLYSVHMKSTKPS--------LSPSKPQSSLVIPTSLFCKTFPFHFMFDKD 293

Qy        227 LVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLD 286
              : : | || | |:: :           | :: | |: :| ||:: :|  ||:| :
Db        294 MTILQFGNGIRRLMNRRDFQGKPNFEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNS 353

Qy        287 VEKLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346
              |:|             : |||||||: |: :|||| || |  |:| | ||||||||:|
Db        354 VKK----------SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIH 403

Qy        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406
              :| ||:||:||| | :  | : |  |    |:   :|||:|||||  || |: |  || :|
Db        404 NALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQL 463

Qy        407 RHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKN 466
                : | ||:: |||:||| |||| | ||: :     :::: :|| ||||||
Db        464 WQGQVVQAKKFSNVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFDQQCGELD- 518

Qy        467 PFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHT 525
              ||||||:    : : || :    ||  |  :|| |||:: :|    || ::: ||:|:
Db        519 --VYKVETIAMPIVWLGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHS 576

Qy        526 GEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL----------MSP 575
              | |  ||:| :|||||||| | | :: |:     |||||  ||| |           |
Db        577 GSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSR 636

Qy        576 ENSDPQFHLEHRG 588
              |   | |  |  |
Db        637 EELPPNFPSEIPG 649


RESULT 14
JH0810
guanylate cyclase (EC 4.6.1.2) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 26-Aug-1999
C;Accession: JH0810
R;Yoshikawa, S.; Miyamoto, I.; Aruga, J.; Furuichi, T.; Okano, H.; Mikoshiba, K.
J. Neurochem. 60, 1570-1573, 1993
A;Title: Isolation of a Drosophila gene encoding a head-specific guanylyl cyclase.
A;Reference number: JH0810; MUID:93203896; PMID:8095978
A;Accession: JH0810
A;Molecule type: mRNA
A;Residues: 1-683 <YOS>
A;Cross-references: GB:S57126; NID:g298674; PID:g298675
A;Experimental source: head
C;Genetics:
A;Gene: dgc 1
```

```
A;Cross-references: FlyBase:FBgn0010227
C;Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C;Keywords: cGMP biosynthesis; phosphorus-oxygen lyase
F;417-641/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             22.0%;  Score 712;  DB 2;  Length 683;
  Best Local Similarity   30.2%;  Pred. No. 2.6e-41;
  Matches  188;  Conservative  121;  Mismatches  229;  Indels   84;  Gaps   20;

Qy         22 WEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVL----NLNAGEILQMFGKMFFVF 77
              | :| |  :|| :       :      || :|   ::|      :| ||| : |:
Db         73 WPNIHK-LKLDPQ-------TFKSCANYDYLADIQELLLKMDEASASEILVLLGEELITC 124

Qy         78 CQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDAEKGKGLILHYYS 137
              |     :   | ||:::::||| :|| ::| |       :    | |     |:| :: : |
Db        125 CCTGIIERAFRCLGTDLQEFLGSLDGVYDVLKLQEEDVTDTGFVCA----GEGELI-FTS 179

Qy        138 EREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC--------------DHTQFLIE 182
              ||  :  :::| :| : : ::  ::::|:  :  | |              ||
Db        180 ERPVIAWLLLGSLKALTRMLYKVDVNIKI--EPVERCPAVSLPLLAGQGQLPDHAD---G 234

Qy        183 EKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL-P 241
              |  :   | : |   :   : :::  :||| ||:| | :  | : | |    ::  |
Db        235 SSTSVSKTIPETVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLYKP 294

Qy        242 QLQPGNCSLLSVFSLVRPH-IDISFHGILSHINTVFV--LRSKEGLLDVEKLECEDELTG 298
              :    |   : |   ||  : : |  |:    | |:  | :  | :|   :
Db        295 YMADFGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIG------- 347

Qy        299 TEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQ 358
                :||||:: ||:::|:||: || :  || ||  ||::|||||||||||:::|:|||
Db        348 -----LEIKGQMVHCPESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQ 402

Qy        359 FREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYD 418
              | :  | : :: : : ::    |:  |:||  :||: : |  :| :|     : || |
Db        403 ARAQDGLRRRMDKIKNSIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYP 462

Qy        419 NVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPF--VYKVETVG 476
              :|||||| |||| : ||:          ::::|  ||  ||   |      |  ||||||:|
Db        463 DVTILFSDIVGFTSICSRAT----PFMVISMLEGLYKDFDEFCD-----FFDVYKVETIG 513

Qy        477 DKYMTVSGLPEPCIHHARSICHLALDMMEI----AGQVQVDGESVQITIGIHTGEVVTGV 532
              | |   |||     |: |     |  || :::    :  :  ||| ::: ||:||| |: ||
Db        514 DAYCVASGLHRASIYDA----HRCLDGLKMIDACSKHITHDGEQIKMRIGLHTGTVLAGV 569

Qy        533 IGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSM 592
              :|::|||||||::| : :: |:  |  |||||  |   |   |    :| |: | | |
Db        570 VGRKMPRYCLFGHSVTIANKFESGSEALKINVSPTTKDWLTKHEGF--EFELQPRDP-SF 626

Qy        593 KGKKEPMQVWFLSRKNTGTEET 614
              |: |          | |  ||
Db        627 LPKEFP---------NPGGTET 639


RESULT 15
F88642
protein C54E4.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: F88642
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: F88642
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-583 <STO>
A;Cross-references: GB:chr_IV; PIDN:AAB92031.1; PID:g2702405; GSPDB:GN00022; CESP:C54E4.
A;Note: Similar to guanylate cyclase
C;Genetics:
A;Gene: C54E4.3
A;Map position: 4

  Query Match             21.7%;  Score 701.5;  DB 2;  Length 583;
  Best Local Similarity   29.6%;  Pred. No. 1.1e-40;
  Matches  172;  Conservative  139;  Mismatches  207;  Indels   63;  Gaps   17;

Qy         26 KKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEILQMFGKMFFVFCQESGYDT 85
              :::|   |: :| ::  |||::|  :   |: || |:  :: :|:|:|       |:|:
Db         20 RRKAGYQEDIKFDIQCYYDDTETMRIFRVAATVLGLSVDDMWEMYGEFLITHACETGWQK 79

Qy         86 ILRVLGSNVREFLQNLDALH---DHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGL 142
              :|  : :|::||| ||:::|   | :|      |: |:|:|    :  || |||:| |:||
Db         80 MLFCMANNLQEFLDNLNSMHYFIDQIA-FKSEMKGPTFQCEPFGE-SGLKLHYFSFRQGL 137

Qy        143 QDIVIGIIKTVAQQIHGTEIDMKV-IQQRNEE------CDHTQFLIEEKESKEED---FY 192
              || |:::  |: :   |:|:|| : :||:|       :|  | :|  ::  :    |:
Db        138 FPIVKGLVRKTARTLF--EMDVKVCMLERNQERRKSGMVEHVIFSVEPDDNHRKGKRLFH 195

Qy        193 EDLDRFEENGTQESRISPYT----------FCKAFPFHIIFDRDLVVTQCGNAIYRVLP 241
              :|    | |:  | :|           |   ||:|: |:: :::   |  : |
Db        196 ----KFRNTKTTENAPS-FTLSSTILVGLRDFKNIFPYHVCFNKQMIIEHIGIYLLREYG 250

Qy        242 QLQPGNCSLLSVFSLVRP-HIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
                :  :  ||:|  | :::  :||::||:|: : | .    :: | ::  :
Db        251 LENKKTLKVSDLMQLVQPSDIQLTYKNVLSYLNTLFIFQLKH---HSKRNEVQEGSSEAF 307

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              | |||:|: : : :||:|:||| |  : |:    ||:||:|:|||||||:| :
Db        308 QQPLVLKGEMMPINDGNSIIFICSPHVTTVRDILNLKLYISDMPMHDATRDLVMLNQSRI 367

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPA------ 414
              : :| ::||    :::      || :| :|| ||:  :|| :|  ||  : |||
Db        368 CQMELNKKLEETMKKMKKMTEELEVKKSQTDRLLFEFVPPVIAEALRAAKTVPALMQKRI 427

Qy        415 -----KRYDNVTILFSGIVGF---NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKN 466
                  : : :::|: |  |   :  ||          :|: :: ||: ||| : :  |
Db        428 SIDFSEEFSDCSVIFTDIPDFFTISVNCS-------PTEIITVVTDLFHRFDRIIEKHKG 480

Qy        467 PFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQVD--GESVQITIGIH 524
              ||| :: | |: | |:|    :|     :||| ::  | || |     ||:: ||:|
Db        481 ---YKVLSLMDSYLIVGGVPNANQYHCEDSLNLALGLLFEAKQVVVPKLERSVRLRIGVH 537

Qy        525 TGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVS 565
              | || |:: |: ||:|: |||||:|    :    ||: ||
Db        538 CGPVVAGIVSQQKPRFCVLGNTVNVTKSICSHSSPGKVLVS 578


Search completed: June 27, 2003, 13:05:25
Job time : 19.4966 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 12:55:07 ; Search time 11.0695 Seconds
(without alignments)
2585.358 Million cell updates/sec

Title: US-09-762-767A-2
Perfect score: 3593
Sequence: 1 MFCTKLKDLKITGECPFSLL.....QKKDVEDGNANFLGKASGID 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3593	100.0	690	1	CYG3_HUMAN	Q02108 homo sapien
2	3232	90.0	690	1	CYG3_RAT	P19686 rattus norv
3	3107.5	86.5	691	1	CYG3_BOVIN	P19687 bos taurus
4	1651	46.0	730	1	CYG4_RAT	Q9wvi4 rattus norv
5	1635	45.5	732	1	CYG4_HUMAN	P33402 homo sapien
6	976.5	27.2	683	1	CYGH_DROME	Q07093 drosophila
7	795	22.1	619	1	CYG1_RAT	P20595 rattus norv
8	793.5	22.1	619	1	CYG1_BOVIN	P16068 bos taurus
9	789.5	22.0	619	1	CYG1_HUMAN	Q02153 homo sapien
10	774.5	21.6	682	1	CYG2_RAT	P22717 rattus norv
11	708	19.7	617	1	CYG2_HUMAN	O75343 homo sapien
12	460	12.8	1047	1	ANPB_BOVIN	P46197 bos taurus
13	460	12.8	1047	1	ANPB_HUMAN	P20594 homo sapien
14	460	12.8	1047	1	ANPB_RAT	P16067 rattus norv
15	457.5	12.7	1057	1	ANPA_RAT	P18910 rattus norv
16	456.5	12.7	1061	1	ANPA_HUMAN	P16066 homo sapien
17	454.5	12.6	1057	1	ANPA_MOUSE	P18293 mus musculu
18	452	12.6	433	1	KSGC_RAT	P55205 rattus norv
19	450	12.5	1108	1	CYGE_MOUSE	P52785 mus musculu
20	448	12.5	1108	1	CYGE_RAT	P51840 rattus norv
21	445	12.4	1109	1	CYGD_CANFA	O19179 canis famil
22	442	12.3	1108	1	CYGF_HUMAN	P51841 homo sapien
23	440	12.2	1103	1	CYGF_BOVIN	O02740 bos taurus
24	438	12.2	1108	1	CYGF_RAT	P51842 rattus norv
25	438	12.2	1110	1	CYGD_BOVIN	P55203 bos taurus
26	434	12.1	1103	1	CYGD_HUMAN	Q02846 homo sapien
27	430	12.0	1110	1	CYGX_RAT	P51839 rattus norv
28	428.5	11.9	1050	1	ANPB_ANGJA	P55202 anguilla ja
29	427	11.9	1125	1	CYGS_STRPU	P16065 strongyloce
30	414.5	11.5	1073	1	HSER_PIG	P55204 sus scrofa
31	408.5	11.4	1073	1	HSER_HUMAN	P25092 homo sapien
32	407.5	11.3	1072	1	HSER_RAT	P23897 rattus norv
33	399.5	11.1	1076	1	HSER_CAVPO	P70106 cavia porce
34	389	10.8	1137	1	CYG1_CAEEL	Q09435 caenorhabdi
35	322.5	9.0	887	1	CYA2_HUMAN	Q08462 homo sapien
36	322	9.0	1064	1	CYA4_RAT	P26770 rattus norv
37	319.5	8.9	1099	1	CYA7_MOUSE	P51829 mus musculu
38	319	8.9	1090	1	CYA2_RAT	P26769 rattus norv
39	316	8.8	1080	1	CYA7_HUMAN	P51828 homo sapien
40	310.5	8.6	1078	1	CYA7_BOVIN	Q29450 bos taurus
41	303.5	8.4	1144	1	CYA3_HUMAN	O60266 homo sapien
42	303	8.4	1144	1	CYA3_RAT	P21932 rattus norv
43	302	8.4	1134	1	CYA1_BOVIN	P19754 bos taurus
44	290.5	8.1	2248	1	CYA1_DROME	P32870 drosophila
45	287.5	8.0	1249	1	CYA8_MOUSE	P97490 mus musculu

ALIGNMENTS

RESULT 1
CYG3_HUMAN
ID CYG3_HUMAN STANDARD; PRT; 690 AA.
AC Q02108; O43843;
DT 01-JUL-1993 (Rel. 26, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)
DE (Soluble guanylate cyclase large subunit) (GCS-alpha-3).
GN GUCY1A1 OR GUCY1A3 OR GUC1A3 OR GUCSA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92316204; PubMed=1352257;
RA Giuili G., Scholl U., Bulle F., Guellaeen G.;
RT "Molecular cloning of the cDNAs coding for the two subunits of
RT soluble guanylyl cyclase from human brain.";
RL FEBS Lett. 304:83-88(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Gansemans Y., Brouckaert P., Fiers W.;
RT "Human soluble guanylate cyclase large subunit mRNA, alpha3-like.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98416113; PubMed=9742212;
RA Zabel U., Weeger M., La M., Schmidt H.H.;
RT "Human soluble guanylate cyclase: functional expression and revised
RT isoenzyme family.";
RL Biochem. J. 335:51-57(1998).
CC -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC ---

CC ---
DR EMBL; X66534; CAA47145.1; -.
DR EMBL; U58855; AAB94794.1; -.

```
DR   EMBL; Y15723; CAA75738.1; -.
DR   PIR; S23098; S23098.
DR   HSSP; P19687; 1AWN.
DR   Genew; HGNC:4685; GUCY1A3.
DR   MIM; 139396; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family.
FT   DOMAIN      481    608       GUANYLATE CYCLASE.
FT   CONFLICT    124    127       AAGV -> QQS (IN REF. 1).
FT   CONFLICT    131    184       VIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLK
FT                                QSSHCQEAGKRGR -> LSKNLLVKRFLKYVTRKMKTSLGW
FT                                LEAPLKIFKQLQYPSETEQPLPRSRKKGQ (IN REF.
FT                                1).
FT   CONFLICT    322    322       MISSING (IN REF. 1).
FT   CONFLICT    529    535       GDAYCVA -> AMPIVWL (IN REF. 1).
FT   CONFLICT    678    690       GNANFLGKASGID -> ASQFFRQSIRNRLATYIPIYKSLG
FT                                FDSLKMCRASESTLGIVDG (IN REF. 1).
SQ   SEQUENCE   690 AA;  77452 MW;  DA1E14A5E11451CF CRC64;

  Query Match           100.0%;  Score 3593;  DB 1;  Length 690;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-243;
  Matches  690;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660

Qy        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690
              |||||||||||||||||||||||||||||||
Db        661 QQGTNSKPCFQKKDVEDGNANFLGKASGID 690


RESULT 2
CYG3_RAT
ID   CYG3_RAT       STANDARD;      PRT;   690 AA.
AC   P19686;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)
DE   (Soluble guanylate cyclase large subunit).
GN   GUCY1A1 OR GUC1A1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=91009100; PubMed=1698769;
RA   Nakane M., Arai K., Saheki S., Kuno T., Buechler W., Murad F.;
RT   "Molecular cloning and expression of cDNAs coding for soluble
RT   guanylate cyclase from rat lung.";
RL   J. Biol. Chem. 265:16841-16845(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Striatum;
RX   MEDLINE=97151525; PubMed=8997507;
RA   Smigrodzki R.M., Levitt P.;
RT   "The alpha 1 subunit of soluble guanylyl cyclase is expressed
RT   prenatally in the rat brain.";
RL   Brain Res. Dev. Brain Res. 97:226-234(1996).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57405; AAA41206.1; -.
DR   EMBL; U60835; AAB17953.1; -.
DR   PIR; A38297; OYRTA1.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family.
FT   DOMAIN      480    607       GUANYLATE CYCLASE.
SQ   SEQUENCE   690 AA;  77566 MW;  E4819B2CA4F86401 CRC64;

  Query Match           90.0%;  Score 3232;  DB 1;  Length 690;
  Best Local Similarity  89.1%;  Pred. No. 4.3e-218;
  Matches  616;  Conservative   39;  Mismatches   34;  Indels    2;  Gaps    2;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLP 60
              ||| | |||||||||||||||||| |  || || ||||:||:| ||:  | | : | |
Db          1 MFCRKFKDLKITGECPFSLLAPGQVPTEPIEEVAGVSESCQATLPTCQEFAE-NAEGSHP 59

Qy         61 QRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIAE 120
```

```
Db         60 QRKTSRNRVYLHTLAESIGKLIFPEFERLNLALQRTLAKHKIKENRNSSEKEDLERIIAE 119

Qy        121 QAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEAG 180

Db        120 EAIAAGVPVEVLKDSLGEELFKICYEEDEHILGVVGGTLKDFLNSFSTLLKQSSHCQEAE 179

Qy        181 KRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCF 240

Db        180 RRGRLEDASILCLDKDQDFLNVYYFFPKRTTALLLPGIIKAAARILYESHVEVSLMPPCF 239

Qy        241 HNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQ 300

Db        240 RSECTEFVNQPYLLYSVHVKSTKPSLSPGKPQSSLVIPTSLFCKTFPFHFMLDRDLAILQ 299

Qy        301 FGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKS 360

Db        300 LGNGIRRLVNKRDFQGKPNFEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRWDNLVKKS 359

Qy        361 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 420

Db        360 SRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQAR 419

Qy        421 AQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNV 480

Db        420 AQDGLKKRLGKLKATLEHAHQALEEEKKKTVDLLCSIFPSEVAQQLWQGQIVQAKKFNEV 479

Qy        481 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHK 540

Db        480 TMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHR 539

Qy        541 ESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600

Db        540 ESDTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 599

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDAY 660

Db        600 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLDAY 659

Qy        661 Q-QGTNSKPCFQKKDVEDGNANFLGKASGID 690

Db        660 QHQGPNSKPWFQQKDAEDGNANFLGKASGVD 690


RESULT 3
CYG3_BOVIN
ID   CYG3_BOVIN     STANDARD;      PRT;   691 AA.
AC   P19687;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)
DE   (Soluble guanylate cyclase large subunit).
GN   GUCY1A1 OR GUC1A1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Adrenal medulla;
RX   MEDLINE=90306336; PubMed=1973124;
RA   Koesling D., Harteneck C., Humbert P., Bosserhoff A., Frank R.,
RA   Schultz G., Boehme E.;
RT   "The primary structure of the larger subunit of soluble guanylyl
RT   cyclase from bovine lung. Homology between the two subunits of the
RT   enzyme.";
RL   FEBS Lett. 266:128-132(1990).
RN   [2]
RP   3D-STRUCTURE MODELING OF 472-628.
RX   MEDLINE=98054247; PubMed=9391039;
RA   Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.;
RT   "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
RT   and mutational analysis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54014; CAA37960.1; -.
DR   PIR; S10713; OYBO77.
DR   PDB; 1AWN; 28-JAN-98.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family; 3D-structure.
FT   DOMAIN      482    609       GUANYLATE CYCLASE.
SQ   SEQUENCE   691 AA;  77532 MW;  5D1FE4D2204E8683 CRC64;

  Query Match             86.5%;  Score 3107.5;  DB 1;  Length 691;
  Best Local Similarity   86.1%;  Pred. No. 2.1e-209;
  Matches  595;  Conservative   41;  Mismatches   54;  Indels    1;  Gaps    1;

Qy          1 MFCTKLKDLKITGECPFSLLAPGQVPNESSEEAAGSSESCKATVP-ICQDIPEKNIQESL 59

Db          1 MFCAKLKDLQITGDCPFSLLAPGQVPREPLGEATGSGPASTPGQPGVCPGVPDKNPPGRL 60

Qy         60 PQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKIKESRKSLEREDFEKTIA 119

Db         61 PRRKTSRSRVYLHTLAESICKLIFPEFERLNLALQRTLAKHKIKENRKSLEREDFEKIVV 120

Qy        120 EQAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLKQSSHCQEA 179

Db        121 DQAIAAGVPVEIIKESLGEELFKICYEEDEYILGVVGGTLKDFLNSFSTLLKQSSHCQEA 180

Qy        180 GKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPC 239

Db        181 EKKGRFEDASILCLDKDPDVLYVYYFFPKRITSLILPGIIKAAARILYETEVEVSSTPSR 240

Qy        240 FHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTIL 299

Db        241 FHQDCREFVDQPCELYSVHIRSARPHPPPGKPVSSLVIPASLFCKTFPFHFMLDRDMSIL 300

Qy        300 QFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK 359

Db        301 QLGHGIRRLMSRRDVQGKPHFDEYFEILTPKISQTFSGIMTMLNMQFLVRVRRWDNSMKK 360

Qy        360 SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQA 419

Db        361 SSRVMDLKGQMIYMVESSSILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQA 420

Qy        420 RAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSN 479

Db        421 RAQDGLKKRLGKLKATLEQAHQALEEEKRKTVDLLCSIFPSEVARQLWQGHAVQAKRFGN 480

Qy        480 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLH 539

Db        481 VTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDRQCGELDVYKVETIGDAYCVAGGLH 540
```

```
Qy        540 KESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 599
              ||||||||||||||||||||| ||:|||||||||||||||||||||||||||||||||||
Db        541 KESDTHAVQIALMALKMMELSHEVVSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGN 600

Qy        600 NVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPGICHFLDA 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||:|
Db        601 NVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLEA 660

Qy        660 YQQGTNSKPCFQKKDVEDGNANFLGKASGID 690
              ||||| ||| |||||||: ||||||||||
Db        661 YQQGTTSKPWFQKKDVEEANANFLGKASGID 691


RESULT 4
CYG4_RAT
ID   CYG4_RAT       STANDARD;      PRT;   730 AA.
AC   Q9WVI4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-2 chain (EC 4.6.1.2) (GCS-alpha-2).
GN   GUCY1A2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar Kyoto; TISSUE=Aorta;
RX   MEDLINE=20571097; PubMed=11121588;
RA   Koglin M., Behrends S.;
RT   "Cloning and functional expression of the rat alpha(2) subunit of
RT   soluble guanylyl cyclase.";
RL   Biochim. Biophys. Acta 1494:286-289(2000).
CC   -!- FUNCTION: HAS GUANYLYL CYCLASE ON BINDING TO THE BETA-1 SUBUNIT.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF109963; AAD42949.2; -.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family.
FT   DOMAIN      519     646       GUANYLATE CYCLASE.
FT   DOMAIN       50      72       ALA-RICH.
FT   DOMAIN       64      72       POLY-ALA.
SQ   SEQUENCE   730 AA;  81786 MW;  B8D790BFF81FB8F9 CRC64;

  Query Match             46.0%;  Score 1651;  DB 1;  Length 730;
  Best Local Similarity   49.2%;  Pred. No. 1.1e-107;
  Matches  355;  Conservative  106;  Mismatches  196;  Indels   64;  Gaps   14;

Qy         13 GECPFSLLA-------PGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESL------ 59
              |||| | |         ||           ||  :  |  |    :|  ::  :
Db         27 GECPLSKLCWNGSRSPPG---------PPGSRAAAMAATP----VPAASVAAAAAAVAAG 73

Qy         60 PQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLA--KHKI---KESRKSL----- 109
               :|   | || | :| |||  |  |  : ::: |:|||   :|::   ::: |:
Db         74 SKRAQRRRRVNLDSLGESISLLTAPSPQTIHMTLKRTLQYYEHQVIGYRDAEKNFHNISN 133

Qy        110 ----------EREDFEKTIAEQAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGT 158
                       | ||    :   |   |:  : |:|  ||| ||||::|:| :|  || |
Db        134 RCSSADHSNKEEIEDVSGILRCTANVLGLKFQEIQERFGEEFFKICFDENERVLRAVGST 193

Qy        159 LKDFLNSFSTLLKQSSHCQEA-GKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPG 217
              |:|| | |  ||:   | : : ||:  ||  | || :  :  | ::|| |  |    : |
Db        194 LQDFFNGFDALLE---HIRTSFGKQATLESPSFLCKELPEGTLKLHYFHPHHTVGFAMLG 250

Qy        218 IIKAAAHVLYETEVEV-SLMPPCFHNDCSEFVNQPYLLYSVHMKSTKPSLSPSKPQSSLV 276
              :||||   :|   |||  :    | :| |   |   | : :    |   :: : || :
Db        251 MIKAAGKRIYHLNVEVEQIENEKFCSDGSTPSNYSCLTFLIKECET-TQITKNIPQGTSQ 309

Qy        277 IPTSL------FCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPK 330
              ||| |      ||:|||| ||| :| :|| | |:|: : | |      ||: |||::||
Db        310 IPTDLRISINTFCRTFPFHLMFDPNMVVLQLGEGLRKQL-RCDNHKVLKFEDCFEIVSPK 368

Qy        331 INQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRL 390
              :| ||  ::  |:  ||:| :   :      :||::|||||:: ||:||||||||||:|
Db        369 VNATFDRVLLRLSTPFVIRTKPEASGTDNEDKVMEIKGQMIHVPESNAILFLGSPCVDKL 428

Qy        391 EDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKT 450
              ::  ||||:|||||||:| |||:|:||||:||||||||: |||||||: ||||||||||
Db        429 DELIGRGLHLSDIPIHDATRDVILVGEQAKAQDGLKKRMDKLKATLEKTHQALEEEKKKT 488

Qy        451 VDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALY 510
              |||| |||| :|||||||| | |||:|| :|||||||||||||||:||:|:|||:||| ||
Db        489 VDLLYSIFPGDVAQQLWQRQQVQARKFDDVTMLFSDIVGFTAICAQCTPMQVISMLNELY 548

Qy        511 TRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEP 570
              |||| ||| ||:||||||||||||| |||::|  ||  |||||||||||:||::| | |
Db        549 TRFDHQCGFLDIYKVETIGDAYCVASGLHRKSLCHAKPIALMALKMMELSEEVLTPDGRP 608

Qy        571 IKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCP 630
              |:||||:||||| ||||||:|||||||||||||:|||| | ||:||:|||||:|||
Db        609 IQMRIGIHSGSVLAGVVGVRMPRYCLFGNNVTLASKFESGSHPRRINISPTTYQLLKRED 668

Qy        631 GFVFTPRSREELPPNFPSEIPGICHFLDAYQQGTNSKPCFQKKDVEDGNAN----FLGKA 686
              | | |||||||| ||| ||||:|:||:        ||      ::  : |    || :
Db        669 SFTFIPRSREELPDNFPKEIPGVCYFLELRTGPKPPKPSLSSSRIKKVSYNIGTMFLRET 728

Qy        687 S 687
              |
Db        729 S 729


RESULT 5
CYG4_HUMAN
ID   CYG4_HUMAN     STANDARD;      PRT;   732 AA.
AC   P33402;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-2 chain (EC 4.6.1.2) (GCS-alpha-2).
GN   GUCY1A2 OR GUC1A2 OR GUCSA2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92070494; PubMed=1683630;
RA   Harteneck C., Wedel B., Koesling D., Malkewitz J., Boehme E.,
RA   Schultz G.;
RT   "Molecular cloning and expression of a new alpha-subunit of soluble
RT   guanylyl cyclase. Interchangeability of the alpha-subunits of the
RT   enzyme.";
RL   FEBS Lett. 292:217-222(1991).
RN   [2]
```

```
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA 2-I).
RX   MEDLINE=95403342; PubMed=7673142;
RA   Behrends S., Harteneck C., Schultz G., Koesling D.;
RT   "A variant of the alpha 2 subunit of soluble guanylyl cyclase
RT   contains an insert homologous to a region within adenylyl cyclases
RT   and functions as a dominant negative protein.";
RL   J. Biol. Chem. 270:21109-21113(1995).
CC   -!- FUNCTION: HAS GUANYLYL CYCLASE ON BINDING TO THE BETA-1 SUBUNIT.
CC       THE ALTERNATIVE SPLICED ISOFORM, ALPHA-2-I, ACTS AS A NEGATIVE
CC       REGULATOR OF GUANYLYL CYLASE ACTIVITY AS IT FORMS NON-FUNCTIONAL
CC       HETERODIMERS WITH THE BETA SUBUNITS.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-2 (SHOWN HERE) AND
CC       ALPHA-2-I; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALPHA-2 ISOFORM IS EXPRESSED IN FETAL BRAIN,
CC       LIVER, COLON, ENDOTHELIUM AND TESTIS, WHILE THE ALPHA-2-I FORM
CC       IS EXPRESSED ONLY IN LIVER, COLON AND ENDOTHELIUM.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63282; CAA44921.1; -.
DR   EMBL; Z50053; CAA90393.1; -.
DR   PIR; S18325; S18325.
DR   HSSP; P19687; 1AWN.
DR   Genew; HGNC:4684; GUCY1A2.
DR   MIM; 601244; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family; Alternative splicing.
FT   DOMAIN      521    648       GUANYLATE CYCLASE.
FT   DOMAIN       51     76       ALA-RICH.
FT   DOMAIN       51     58       POLY-ALA.
FT   VARSPLIC    612    612       Q -> QPQRSELLFSFPVSIQLVPDQHQSETDLGTEK
FT                                (IN ISOFORM ALPHA 2-I).
SQ   SEQUENCE   732 AA;  81749 MW;  79465A7D3FE52DB7 CRC64;

  Query Match           45.5%;  Score 1635;  DB 1;  Length 732;
  Best Local Similarity   49.4%;  Pred. No. 1.5e-106;
  Matches  355;  Conservative  106;  Mismatches  200;  Indels   58;  Gaps   15;

Qy         13 GECPFSLLA-------PGQVPNESSEEAAGSSESCKATVPICQDIPEKNIQESLPQRKTS 65
              |||| | |        ||  | | |  || ::    |  |          :  :|
Db         27 GECPLSRLCWNGSRSPPG--PLEPSPAAAAAA---AAPAPTPAASAAAAAATAGARRVQR 81

Qy         66 RSRVYLHTLAESICKLIFPEFERLNVALQRTLA--KHKI---KESRKSL---------- 109
              | || | :| ||| :|  |  : :   |:|||   :|::   ::: |:
Db         82 RRRVNLDSLGESISRLTAPSPQTIQQTLKRTLQYYEHQVIGYRDAEKNFHNISNRCSYAD 141

Qy        110 -----EREDFEKTIAEQAVAAGVPVEVIKESLGEEVFKICYEEDENILGVVGGTLKDFLN 164
                  | ||    :   |   |:  | |::  ||| | ||: |:| :|  |||||:|| |
Db        142 HSNKEEIEDVSGILQCTANILGLKFEEIQKRFGEEFFNICFHENERVLRAVGGTLQDFFN 201

Qy        165 SFSTLLKQSSHCQEA-GKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAA 223
               |  ||:   | : : ||:  ||  | || :  :  | ::|| |       : |:|||
Db        202 GFDALLE---HIRTSFGKQATLESPSFLCKELPEGTLMLHYFHPHHIVGFAMLGMIKAAG 258

Qy        224 HVLYETEVEVSLMPPCFHND--CSEFVNQP---YLLYSVHMKSTKPSLSPSKPQSSLVIP 278
               :|  :|||  :     |:  ||: |: |     |  :  :    ::  : || :  :|
Db        259 KKIYRLDVEVEQVA----NEKLCSD-VSNPGNCSCLTFLIKECENTNIMKNLPQGTSQVP 313

Qy        279 TSL------FCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKIN 332
               |      ||: |||| |||  |::|| | |:|: : | |      ||: |||::||:|
Db        314 ADLRISINTFCRAFPFHLMFDPSMSVLQLGEGLRKQL-RCDTHKVLKFEDCFEIVSPKVN 372

Qy        333 QTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIYIVESSAILFLGSPCVDRLED 392
               ||  ::  |:  ||:| :   :  :   :||::|||||:: ||::||||||||||:|::
Db        373 ATFERVLLRLSTPFVIRTKPEASGSENKDKVMEVKGQMIHVPESNSILFLGSPCVDKLDE 432

Qy        393 FTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVD 452
               ||||:|||||||:| |||:|:||||:||||||||: |||||||: ||||||||||||
Db        433 LMGRGLHLSDIPIHDATRDVILVGEQAKAQDGLKKRMDKLKATLERTHQALEEEKKKTVD 492

Qy        453 LLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNALYTR 512
               || |||| :|||||||| |||:|| :|||||||||||||||:||:|:|||:||| ||||
Db        493 LLYSIFPGDVAQQLWQGQQVQARKFDDVTMLFSDIVGFTAICAQCTPMQVISMLNELYTR 552

Qy        513 FDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIK 572
               || ||| ||:|||||||||||| |||::|  ||  ||||||||||||:|||::| | ||:
Db        553 FDHQCGFLDIYKVETIGDAYCVAAGLHRKSLCHAKPIALMALKMMELSEEVLTPDGRPIQ 612

Qy        573 MRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGF 632
               ||||:|||| ||||||:|||||||||||||:|||| | ||:|||||||:|||     |
Db        613 MRIGIHSGSVLAGVVGVRMPRYCLFGNNVTLASKFESGSHPRRINVSPTTYQLLKREESF 672

Qy        633 VFTPRSREELPPNFPSEIPGICHFLDAYQQGTNSKPCFQKKDVEDGNAN----FLGKAS 687
               | |||||||| ||| ||||||:||:        ||      ::  : |    || : |
Db        673 TFIPRSREELPDNFPKEIPGICYFLEVRTGPKPPKPSLSSSRIKKVSYNIGTMFLRETS 731


RESULT 6
CYGH_DROME
ID   CYGH_DROME     STANDARD;      PRT;   683 AA.
AC   Q07093;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Head-specific guanylate cyclase (EC 4.6.1.2).
GN   GYC-ALPHA-99B OR GYC-ALPHA-63A OR GYC OR DGC1.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Canton-S; TISSUE=Head;
RX   MEDLINE=93203896; PubMed=8095978;
RA   Yoshikawa S., Miyamoto I., Aruga J., Furuichi T., Okano H.,
RA   Mikoshiba K.;
RT   "Isolation of a Drosophila gene encoding a head-specific guanylyl
RT   cyclase.";
RL   J. Neurochem. 60:1570-1573(1993).
CC   -!- FUNCTION: MAY HAVE A ROLE IN PHOTOTRANSDUCTION. A SECOND SUBUNIT
CC       MAY BE REQUIRED FOR ENZYME ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBUNIT: DIMER (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- TISSUE SPECIFICITY: HEAD, WHERE IT IS PREFERENTIALLY EXPRESSED IN
CC       THE CNS AND THE RETINA. NOT FOUND IN BODIES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; S57126; AAB25820.1; -.
DR   PIR; JH0810; JH0810.
DR   HSSP; P19687; 1AWN.
DR   FlyBase; FBgn0013972; Gyc-alpha-99B.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family; Vision.
FT   DOMAIN      465    591       GUANYLATE CYCLASE.
SQ   SEQUENCE   683 AA;  75906 MW;  12D67C4B79A07C8D CRC64;

  Query Match           27.2%;  Score 976.5;  DB 1;  Length 683;
  Best Local Similarity  37.3%;  Pred. No. 1.3e-60;
  Matches  244;  Conservative  116;  Mismatches  211;  Indels   83;  Gaps   20;

Qy         58 SLPQRKTSRSRVYLHTLAESICKLIFPEFERLNVALQRTLAK--------HKIK---ESR 106
              :|·  : |   : |  |  :|  |  |  | .|| |:   :||        ||:|   ::
Db         27 ALEDEELSDDALTLTHLQMAIQLLTAPSNEDLNTAVTSLVAKYRQNWPNIHKLKLDPQTF 86

Qy        107 KSLEREDFEKTIAE------QAVAAGVPVEVIKESLGEEVFKICYEE-DENILGVVGGTL 159
              ||    |:   | |      :| |: : |      ||||:   |      |     :|  |
Db         87 KSCANYDYLADIQELLLKMDEASASEILV-----LLGEELITCCCTGIIERAFRCLGTDL 141

Qy        160 KDFLNSFS---TLLKQSSHCQEAGKRGRLEDASILCLDKEDDFLHVYYFFPKRTTSLILP 216
              ::|| |      :||      ||      : |   :|  : :   : :   :  : :|
Db        142 QEFLGSLDGVYDVLK----LQEED----VTDTGFVCAGEGE----LIFTSERPVIAWLLL 189

Qy        217 GIIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQPYLL----YSVHMKSTKPSLSPSKPQ 272
              | :||   :||: :| : : |     :    |: | |        |   :  |:| : |:
Db        190 GSLKALTRMLYKVDVNIKIEPV----ERCPAVSLPLLAGQGQLPDHADGSSTSVSKTIPE 245

Qy        273 ----------SSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNFEE 322
                       | | : :| ||| ||:||: :: : ::| | |  :|        ||   :
Db        246 TVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLY-------KPYMAD 298

Qy        323 -------YFEILTPK-INQTFSGIMTMLNMQFVVRVRRWDNSVKKSSRVMDLKGQMIYIV 374
                    ||:   || :   |  |:       |:: :     :|   :  :::||||::
Db        299 FGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIGLEIKGQMVHCP 358

Qy        375 ESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGKLKA 434
              ||:::||:||| :| |:  |  ||::||||:|:| |:|:|:||||||||||::|: |:|
Db        359 ESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQARAQDGLRRRMDKIKN 418

Qy        435 TLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAIC 494
              ::|:|: |: :|:|| | ||  ||| |:|::|| |  : || : :||:|||||||||:||
Db        419 SIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYPDVTILFSDIVGFTSIC 478

Qy        495 SQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKES--DTHAVQIALM 552
              |: :|  ||:||  ||  ||: |  ||||||||||||||| |||: |  | |      |
Db        479 SRATPFMVISMLEGLYKDFDEFCDFFDVYKVETIGDAYCVASGLHRASIYDAHR---CLD 535

Qy        553 ALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
               |||::   : ::  || |||||||:|:| ||||| ||||||||::||:|||||| |
Db        536 GLKMIDACSKHITHDGEQIKMRIGLHTGTVLAGVVGRKMPRYCLFGHSVTIANKFESGSE 595

Qy        613 PRKINVSPTTYRLLKDCPGFVF--TPRSREELPPNFPSEIPG---ICHFLDAYQ 661
              ||||||||   |    || |   ||   ||  ||:  ||    |:||::::
Db        596 ALKINVSPTTKDWLTKHEGFEFELQPRDPSFLPKEFPN--PGGTETCYFLESFR 647


RESULT 7
CYG1_RAT
ID   CYG1_RAT       STANDARD;      PRT;   619 AA.
AC   P20595;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
DE   (Soluble guanylate cyclase small subunit).
GN   GUCY1B1 OR GUCY1B3 OR GUC1B3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89087429; PubMed=2905128;
RA   Nakane M., Saheki S., Kuno T., Ishii K., Murad F.;
RT   "Molecular cloning of a cDNA coding for 70 kilodalton subunit of
RT   soluble guanylate cyclase from rat lung.";
RL   Biochem. Biophys. Res. Commun. 157:1139-1147(1988).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: LUNG AND BRAIN.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M22562; AAA41204.1; -.
DR   PIR; A31871; OYRTB1.
DR   HSSP; P16068; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis.
FT   DOMAIN      421    554       GUANYLATE CYCLASE.
SQ   SEQUENCE   619 AA;  70477 MW;  BCC9850DE4504226 CRC64;

  Query Match           22.1%;  Score 795;  DB 1;  Length 619;
  Best Local Similarity  34.3%;  Pred. No. 5.2e-48;
  Matches  210;  Conservative  100;  Mismatches  225;  Indels   78;  Gaps   15;

Qy         89 LNVALQRTLAKHKIKESRKSLERE--------------DFEKT---IAEQAVAAGVPVE 130
              :| ||:  : ::   |  : :::|              |  ||   :|  :    :
Db          5 VNHALELLVIRNYGPEVWEDIKKEAQLDQEGQFLVRIIYDHSKTYDLVAAASKVLNLNAG 64

Qy        131 VIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDAS 189
              | :  |:  |  | |   : || |:|  :::|| :   |     |  |     :  |
Db         65 EILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDAL-----HDHLATIYPGMRAPS 119

Qy        190 ILCLDKE-DDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLM----PPCFHNDC 244
              | | |    | ::|:  :     |: ||||  |  :: ||::: ::      | |
Db        120 FRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRSEECDHTQ- 178

Qy        245 SEFVNQPYLLYSVHMKSTKPS--------LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDM 296
                    : :  | :|             :  | | : | : ||| |||| :||:|:
Db        179 ----------FLIEEKESKEEDFYEDLDRFEENGTQDSRISPYT-FCKAFPFHIIFDRDL 227

Qy        297 TILQFGNGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNS 356
              : | || | |:: :    || :    | :: | |: :| ||:: :|  ||:| :
Db        228 VVTQCGNAIYRVLPQLQ-PGKCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLD 286

Qy        357 VKK----------SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIH 406
              |:|             : |||||||: |: :|||| || |  |:| | |||||||||:|
Db        287 VEKLECEDELTGAEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346

Qy        407 NALRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQL 466
              :| ||:||:||| | :  | : |  |    |:   :|||:|||||  || |: |  || :|
Db        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406
```

```
Qy        467 WQGQVVQAKKFSNVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFDQQCGELD- 521
                  : | ||:: |||:||| |||| | ||: :     :::: :|| ||||||
Db        407 RHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKN 466

Qy        522 --VYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHS 579
                |||||:|| |    || :    ||  |  :|| |||:: :|    || ::: ||:|:
Db        467 PFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHT 525

Qy        580 GSVFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSR 639
              | |  ||:| :||||||||| | | :: |:     |||||  ||| |           |
Db        526 GEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL----------MSP 575

Qy        640 EELPPNFPSEIPG 652
              |   | |  |  |
Db        576 ENSDPQFHLEHRG 588


RESULT 8
CYG1_BOVIN
ID   CYG1_BOVIN     STANDARD;      PRT;   619 AA.
AC   P16068;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
DE   (Soluble guanylate cyclase small subunit).
GN   GUCY1B1 OR GUCY1B3 OR GUC1B3.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=89031214; PubMed=2903071;
RA   Koesling D., Herz J., Gausepohl H., Niroomand F., Hinsch K.-D.,
RA   Muelsch A., Boehme E., Schultz G., Frank R.;
RT   "The primary structure of the 70 kDa subunit of bovine soluble
RT   guanylate cyclase.";
RL   FEBS Lett. 239:29-34(1988).
RN   [2]
RP   3D-STRUCTURE MODELING OF 412-572.
RX   MEDLINE=98054247; PubMed=9391039;
RA   Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.;
RT   "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
RT   and mutational analysis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: LUNG AND BRAIN.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00770; CAA68739.1; -.
DR   PIR; S01653; OYBO70.
DR   PDB; 1AWN; 28-JAN-98.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; 3D-structure.
FT   DOMAIN      421    554       GUANYLATE CYCLASE.
SQ   SEQUENCE   619 AA;  70502 MW;  8EFB14952B80F344 CRC64;

  Query Match             22.1%;  Score 793.5;  DB 1;  Length 619;
  Best Local Similarity   35.3%;  Pred. No. 6.6e-48;
  Matches  208;  Conservative   97;  Mismatches  217;  Indels   67;  Gaps   16;

Qy        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153
              :: |  : | |  :: :     |||   |      |:: :  |:  |  | |   : ||
Db         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88

Qy        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212
              |:| ' :::|| :   |     |   |      :   |  | | :    | ::|:  :
Db         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDADKGKGLILHYYSEREGLQ 143

Qy        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265
              |: ||||  |  :: ||::: ::      | |             : :  | :|
Db        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192

Qy        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
                       :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250

Qy        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKSS----------RVMDLKGQM 370
                | :: | |: :| ||:: :|  ||:| :     |:||             : |||||
Db        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKSECEDELTGTEISCLRLKGQM 310

Qy        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430
              ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370

Qy        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490
              |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430

Qy        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543
              | ||: :     :::: :|| ||||||          ||||||:|| |    || :
Db        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490

Qy        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603
              ||  |  :|| |||:: :|    || ::: ||:|:|  |  ||:| :||||||||| | |
Db        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549

Qy        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652
              :: |:     |||||  ||| |         ||   |   | |  | |
Db        550 TSRTETTGEKGKINVSEYTYRCL-------MTP---ENSDPQFHLEHRG 588


RESULT 9
CYG1_HUMAN
ID   CYG1_HUMAN     STANDARD;      PRT;   619 AA.
AC   Q02153;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
DE   (Soluble guanylate cyclase small subunit) (GCS-beta-3).
GN   GUCY1B1 OR GUCY1B3 OR GUC1B3 OR GUCSB3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=92316204; PubMed=1352257;
RA   Giuili G., Scholl U., Bulle F., Guellaeen G.;
RT   "Molecular cloning of the cDNAs coding for the two subunits of
RT   soluble guanylyl cyclase from human brain.";
```

```
RL   FEBS Lett. 304:83-88(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM HSGC-2).
RC   TISSUE=Kidney;
RA   Gansemans Y., Brouckaert P., Fiers W.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 337-545, AND ALTERNATIVE SPLICING.
RC   TISSUE=Lung;
RX   MEDLINE=92008652; PubMed=1680753;
RA   Chhajilani V., Fraendberg P.-A., Ahlner J., Axelsson K.L.,
RA   Wikberg J.E.S.;
RT   "Heterogeneity in human soluble guanylate cyclase due to alternative
RT   splicing.";
RL   FEBS Lett. 290:157-158(1991).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HSGC-1 (SHOWN HERE) AND HSGC-2;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66533; CAA47144.1; -.
DR   EMBL; AF020340; AAB94877.1; -.
DR   PIR; S23097; S23097.
DR   HSSP; P16068; 1AWN.
DR   Genew; HGNC:4687; GUCY1B3.
DR   MIM; 139397; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Alternative splicing.
FT   DOMAIN      421    554       GUANYLATE CYCLASE.
FT   VARSPLIC    393    425       MISSING (IN ISOFORM HSGC-2).
SQ   SEQUENCE   619 AA;  70514 MW;  231E4E660DE02AA1 CRC64;

  Query Match           22.0%;  Score 789.5;  DB 1;  Length 619;
  Best Local Similarity  35.1%;  Pred. No. 1.3e-47;
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy       101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153
                 ::  :  : :  :  ::  :     |||    |      |:: :  |:  |  | |   : ||
Db        30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88

Qy       154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212
                |:|  :::|| :   |     |   |      :   |  | | |    | ::|:  :
Db        89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143

Qy       213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265
                 |: ||||  |  :: ||::: ::      | |             : :  | :|
Db       144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ----------FLIEEKESKEEDFY 192

Qy       266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320
                  :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db       193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250

Qy       321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370
                 | :: | |: :| ||:: :|  ||:| :     |:|             : |||||

Db       251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310

Qy       371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430
                 ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db       311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370

Qy       431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490
                  |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db       371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430

Qy       491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543
                 | ||: :     :::: :|| ||||||          ||||||:|| |    || :
Db       431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490

Qy       544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603
                 ||  |  :|| |||:: :|    || ::: ||:|:|  |  ||:| :|||||||| | |
Db       491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549

Qy       604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652
                 :: |:     |||||  ||| |           | |   | |  | |
Db       550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588


RESULT 10
CYG2_RAT
ID   CYG2_RAT       STANDARD;      PRT;   682 AA.
AC   P22717;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-2 chain (EC 4.6.1.2) (GCS-beta-2).
GN   GUCY1B2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=91105012; PubMed=1980215;
RA   Yuen P.S.T., Potter L.R., Garbers D.L.;
RT   "A new form of guanylyl cyclase is preferentially expressed in rat
RT   kidney.";
RL   Biochemistry 29:10872-10878(1990).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: KIDNEY AND LIVER.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57507; AAA41207.1; -.
DR   PIR; A36228; OYRTB2.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis.
```

```
FT   DOMAIN      408    536       GUANYLATE CYCLASE.
SQ   SEQUENCE   682 AA;  76196 MW;  98C173C1A1CC7715 CRC64;

  Query Match             21.6%;  Score 774.5;  DB 1;  Length 682;
  Best Local Similarity   32.6%;  Pred. No. 1.6e-46;
  Matches  190;  Conservative  107;  Mismatches  204;  Indels   81;  Gaps   10;

Qy        129 VEVIKESLGEEVFKICYEED-ENILGVVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLED 187
              :| | :  ||  || |     : :|  :|| | :|: :   |     |   |     :
Db          1 MEAILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDAL-----HSYLALSYQEMNA 55

Qy        188 ASILCLDKEDDFLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCFHNDCSEF 247
                |    :  |  : ::|:  :     |:||||:| |   ::|:| :|::      | :|
Db         56 PSFRVEEGADGAMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMSIL------DMNEE 109

Qy        248 V----NQPYLLYSVHMKSTKP--SLSPSKPQSS---------------------------- 274
              |     : :::: |  |: :      |:|| |
Db        110 VERTGKKEHVVFLVVQKAHRQIRGAKASRPQGSEDSQADQEALQGTLLRMKERYLNIPVC 169

Qy        275 -------------------------------LVIPTSLFCKTFPFHFMFDKDMTILQFG 302
                                            | :   :||  |||| :||: : : | |
Db        170 PGEKSHSTAVRASVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFHIVFDEALRVKQAG 229

Qy        303 NGIRRLMNRRDFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRR-WDNSVKKSS 361
                |:: :     | |   :||| |: |::    | |   :| |||:: |:       :||
Db        230 VNIQKYVPGILTQ-KFALDEYFSIIHPQVTFNISSICKFINSQFVLKTRKEMMPKARKSQ 288

Qy        362 RVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARA 421
              :: |:||||::     ::|: || |  |::     ::|||| |:  ||::|: :|  |
Db        289 PMLKLRGQMIWMESLRCMIFMCSPNVRSLQELEESKMHLSDIAPHDTTRDLILLNQQRLA 348

Qy        422 QDGLKKRLGKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVT 481
              :  |  :| | |  |       |  |||||  || :: |  || || :|: | | :|   |
Db        349 EMELSCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVANQLKEGRKVAAGEFETCT 408

Qy        482 MLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKE 541
              :||||:| || ||: | |:|:: |||::|::||:     ||||||||||| | ||:
Db        409 ILFSDVVTFTNICAACEPIQIVNMLNSMYSKFDRLTSVHDVYKVETIGDAYMVVGGVPVP 468

Qy        542 SDTHAVQIALMALKMMELSDEVMSP-HGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNN 600
              ::|| ::|  || |   : |||:|  ||||::|:|:|:| | ||||| |||||||||:
Db        469 VESHAQRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVVGDKMPRYCLFGDT 528

Qy        601 VTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              |  |:: ||  :| |:::||| :| ||:  ||    |   |:
Db        529 VNTASRMESHGLPSKVHLSPTAHRALKN-KGFEIVRRGEIEV 569


RESULT 11
CYG2_HUMAN
ID   CYG2_HUMAN     STANDARD;      PRT;   617 AA.
AC   O75343; Q9NZ64;
DT   15-JUL-1999 (Rel. 38, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-2 chain (EC 4.6.1.2) (GCS-beta-2).
GN   GUCY1B2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND REVISION TO 175.
RC   TISSUE=Kidney;
RX   MEDLINE=20241821; PubMed=10777682;
RA   Behrends S., Vehse K.;
RT   "The beta(2) subunit of soluble guanylyl cyclase contains a
RT   human-specific frameshift and is expressed in gastric carcinoma.";
RL   Biochem. Biophys. Res. Commun. 271:64-69(2000).
RN   [2]
RP   SEQUENCE OF 95-257 FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=99107820; PubMed=9889008;
RA   Behrends S., Kazmierczak B., Steenpa A., Knauf B., Bullerdiek J.,
RA   Scholz H., Eiberg H.;
RT   "Assignment of GUCY1B2, the gene coding for the beta2 subunit of human
RT   guanylyl cyclase to chromosomal band 13q14.3 between markers D13S168
RT   and D13S155.";
RL   Genomics 55:126-127(1999).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF038499; AAD09440.2; -.
DR   EMBL; AF218383; AAF66105.1; -.
DR   HSSP; P19687; 1AWN.
DR   Genew; HGNC:4686; GUCY1B2.
DR   MIM; 603695; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis.
FT   DOMAIN      391    519       GUANYLATE CYCLASE.
SQ   SEQUENCE   617 AA;  70367 MW;  1B6C965CAB9A5E3E CRC64;

  Query Match             19.7%;  Score 708;  DB 1;  Length 617;
  Best Local Similarity   31.3%;  Pred. No. 6.1e-42;
  Matches  179;  Conservative  100;  Mismatches  191;  Indels  102;  Gaps   12;

Qy        149 ENILGVVGGTLKDFLNSFSTL----------LKQSSHCQEAGKRGRLEDASILCLDKEDD 198
              : :|  :|| | :|: :   |              :   |   | |  |::
Db          5 DRMLRTLGGNLMEFIENLDALHSYLALSYQEMNAPSFRVERGADGKM---------- 51

Qy        199 FLHVYYFFPKRTTSLILPGIIKAAAHVLYETEVEVSLMPPCFHNDCSEFVNQP------- 251
              |||  |:  :     |:||||:| |   :: :| : ::       | :| | :
Db         52 FLH--YYSDRSGLCHIVPGIIEAVAKDFFDIDVIMDIL------DMNEEVERTGKKEHVV 103

Qy        252 -YLLYSVH--MKSTKP---------------------------------------------- 264
               ::   |  |: |||
Db        104 FLIVQKAHRKMRKTKPKRLQDSQGMERDQEALQAAFLKMKEKYLNVSACPVKKSHWDVVR 163

Qy        265 ------------SLSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRR 312
                         :  |  |:  | |    ||  |||| :||: : : |    |:: :
Db        164 SIVMFGKGHLMNTFEPIYPE-RLWIEEKTFCNAFPFHIVFDESLQVKQARVNIQKYVPGL 222

Qy        313 DFQGKPNFEEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKKSSR-VMDLKGQMI 371
                |     :||| |: |::      |   :| |||:: ||    |   ||  : |:||||
Db        223 QTQ-NIQLDEYFSIIHPQVTFNIFSIRRFINSQFVLKTRREMMPVAWQSRTTLKLQGQMI 281

Qy        372 YIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLGK 431
              ::     :::| || :  |::      ::|||  ::  ||:::|: :|  |:  |  :| :
Db        282 WMESMWCMVYLCSPKLRSLQELEELNMHLSDIAPNDTTRDLILLNQQRLAEIELSNQLER 341

Qy        432 LKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFT 491
              | |:   : |  |||||  || :: |  || || :|: | | :| : |:||||:| ||
Db        342 KKEELQVLSKHLAIEKKKTETLLYAMLPKHVANQLREGKKVAAGEFKSCTILFSDVVTFT 401

Qy        492 AICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGLHKESDTHAVQIAL 551
              ||: | |:|:: :||::|::||:      |||||||||||| | ||:      || ::|
Db        402 NICTACEPIQIVNVLNSMYSKFDRLTSVHAVYKVETIGDAYMVVGGVPVPIGNHAQRVAN 461
```

```
Qy        552 MALKMMELSDEVMSP-HGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESC 610
              || |   : || :|  ||||::|:|:|:| | | ||| |||||||||: |  |:: ||
Db        462 FALGMRISAKEVTNPVTGEPIQLRVGIHTGPVLADVVGDKMPRYCLFGDTVNTASRMESH 521

Qy        611 SVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              :| |:::||| || ||:  ||    |   |:
Db        522 GLPNKVHLSPTAYRALKN-QGFKIIERGEIEV 552


RESULT 12
ANPB_BOVIN
ID   ANPB_BOVIN     STANDARD;      PRT;  1047 AA.
AC   P46197;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B)
DE   (Guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide
DE   B-type receptor).
GN   NPR2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=95147885; PubMed=7845391;
RA   Fenrick R., Babinski K., McNicoll N., Therrien N., Drouin J.,
RA   de Lean A.;
RT   "Cloning and functional expression of the bovine natriuretic peptide
RT   receptor-B (natriuretic factor R1c subtype.";
RL   Mol. Cell. Biochem. 137:173-182(1994).
RN   [2]
RP   SEQUENCE OF 551-1047 FROM N.A.
RC   TISSUE=Retina;
RX   MEDLINE=93371497; PubMed=8103329;
RA   Shmukler B.E., Zubov D.V., Abdulaev N.G.;
RT   "Detection of expression of a membrane form of the guanylate cyclase
RT   type of GC-B in cattle retina (letter).";
RL   Bioorg. Khim. 19:682-685(1993).
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF LIGAND. THE ACTIVATION ORDER SEEMS
CC       TO BE CNP > BNP > ANF.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L26359; AAC41619.1; -.
DR   EMBL; X66865; CAA47334.1; -.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001170; Ntpep_receptorN.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23   1047       ATRIAL NATRIURETIC PEPTIDE RECEPTOR B.
FT   DOMAIN      23    458       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   459    478       POTENTIAL.
FT   DOMAIN     479   1047       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     513    786       PROTEIN KINASE-LIKE.
FT   DOMAIN     861    991       GUANYLATE CYCLASE.
FT   DISULFID    75    101       BY SIMILARITY.
FT   DISULFID   439    439       INTERCHAIN (PROBABLE).
FT   DISULFID   448    448       INTERCHAIN (PROBABLE).
FT   CARBOHYD    24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   958    958       I -> L (IN REF. 2).
SQ   SEQUENCE  1047 AA;  117145 MW;  1A3814D02F22D64F CRC64;

  Query Match             12.8%;  Score 460;  DB 1;  Length 1047;
  Best Local Similarity   41.9%;  Pred. No. 2.6e-24;
  Matches  113;  Conservative   39;  Mismatches   76;  Indels   42;  Gaps    6;

Qy        411 DVVLIGEQARAQD-----------GLKKRLGK--------------------LKATLEQA 439
              ::||: |:  |||           |  :|  |                    |:  :|:
Db        760 ELVLLMERCWAQDPAERPDFGQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEER 819

Qy        440 HQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSP 499
              ||  |||:|   ||  | |  ||:|| :|: |||: | :||: |||||||||: :: :|
Db        820 TQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAESTP 879

Qy        500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL-HKESDTHAVQIALMALKMME 558
              :||:|:|| ||| ||       |||||||||||| |  ||  :    || :|| ||| ::
Db        880 MQVVTLLNDLYTCFDAIIDNFDVYKVETIGDAYMVVSGLPGRNGQRHAPEIARMALALL- 938

Qy        559 LSDEVMS------PHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
              | | |      || : :::|||:|:| | |||||:|||||||||: |  |:: ||
Db        939 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGQ 995

Qy        613 PRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              ||:|| ||   | :   |    |   |:
Db        996 ALKIHVSSTTKDALDELGCFQLELRGDVEM 1025


RESULT 13
ANPB_HUMAN
ID   ANPB_HUMAN     STANDARD;      PRT;  1047 AA.
AC   P20594; Q9UQ50; O60871;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B)
DE   (Guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide
DE   B-type receptor).
GN   NPR2 OR ANPRB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
```

```
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   TISSUE=Brain;
RX   MEDLINE=89365195; PubMed=2570358;
RA   Chang M.S., Lowe D.G., Lewis M., Hellmiss R., Chen E., Goeddel D.V.;
RT   "Differential activation by atrial and brain natriuretic peptides of
RT   two different receptor guanylate cyclases.";
RL   Nature 341:68-72(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   TISSUE=Blood;
RX   MEDLINE=99182382; PubMed=10082481;
RA   Rehemudula D., Nakayama T., Soma M., Takahashi Y., Uwabo J., Sato M.,
RA   Izumi Y., Kanmatsuse K., Ozawa Y.;
RT   "Structure of the type B human natriuretic peptide receptor gene and
RT   association of a novel microsatellite polymorphism with essential
RT   hypertension.";
RL   Circ. Res. 84:605-610(1999).
RN   [3]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Kidney;
RA   Hirsch J.R., Meyer M., Magert H.J., Forssmann W.G., Mollerup S.,
RA   Herter P., Weber G., Cermak R., Ankorina-Stark I., Schlatter E.,
RA   Kruhoffer M.;
RT   "cGMP-dependent and independent inhibition of a K+ conductance by
RT   natriuretic peptides. Molecular and functional studies in human
RT   proximal tubule cells.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF. SEEMS TO BE STIMULATED MORE
CC       EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM/NPR-BI; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB005647; BAA81737.1; -.
DR   EMBL; AB005626; BAA81737.1; JOINED.
DR   EMBL; AB005627; BAA81737.1; JOINED.
DR   EMBL; AB005628; BAA81737.1; JOINED.
DR   EMBL; AB005629; BAA81737.1; JOINED.
DR   EMBL; AB005630; BAA81737.1; JOINED.
DR   EMBL; AB005631; BAA81737.1; JOINED.
DR   EMBL; AB005632; BAA81737.1; JOINED.
DR   EMBL; AB005633; BAA81737.1; JOINED.
DR   EMBL; AB005634; BAA81737.1; JOINED.
DR   EMBL; AB005635; BAA81737.1; JOINED.
DR   EMBL; AB005636; BAA81737.1; JOINED.
DR   EMBL; AB005637; BAA81737.1; JOINED.
DR   EMBL; AB005638; BAA81737.1; JOINED.
DR   EMBL; AB005639; BAA81737.1; JOINED.
DR   EMBL; AB005640; BAA81737.1; JOINED.
DR   EMBL; AB005641; BAA81737.1; JOINED.
DR   EMBL; AB005642; BAA81737.1; JOINED.
DR   EMBL; AB005643; BAA81737.1; JOINED.
DR   EMBL; AB005644; BAA81737.1; JOINED.
DR   EMBL; AB005645; BAA81737.1; JOINED.
DR   EMBL; AB005646; BAA81737.1; JOINED.
DR   EMBL; AJ005282; CAA06466.1; -.
DR   PIR; S05514; S05514.
DR   HSSP; Q02846; 1AWL.
DR   Genew; HGNC:7944; NPR2.
DR   MIM; 108961; -.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001170; Ntpep_receptorN.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal; Alternative splicing; Polymorphism.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23   1047       ATRIAL NATRIURETIC PEPTIDE RECEPTOR B.
FT   DOMAIN      23    458       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   459    478       POTENTIAL.
FT   DOMAIN     479   1047       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     513    786       PROTEIN KINASE-LIKE.
FT   DOMAIN     861    991       GUANYLATE CYCLASE.
FT   DISULFID    75    101       BY SIMILARITY.
FT   DISULFID   439    439       INTERCHAIN (PROBABLE).
FT   DISULFID   448    448       INTERCHAIN (PROBABLE).
FT   CARBOHYD    24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   964   1047       PVCAGVVGLKMPRYCLFGDTVNTASRMESNGQALKIHVSST
FT                               TKDALDELGCFQLELRGDVEMKGKGKMRTYWLLGERKGPPG
FT                               LL -> KADSHSSPSLHLSQTLPTCFFSKGQSVLGLLA
FT                               (IN SHORT ISOFORM).
FT   VARIANT    771    771       Q -> E (IN DBSNP:5816).
FT                               /FTId=VAR_011968.
FT   CONFLICT   755    755       T -> S (IN REF. 2).
SQ   SEQUENCE   1047 AA;  117021 MW;  817FB74D6B31F7EF CRC64;

  Query Match           12.8%;  Score 460;  DB 1;  Length 1047;
  Best Local Similarity  41.9%;  Pred. No. 2.6e-24;
  Matches  113;  Conservative   39;  Mismatches   76;  Indels   42;  Gaps    6;

Qy        411 DVVLIGEQARAQD----------GLKKRLGK-------------------LKATLEQA 439
              ::||: |:  |||           |  :|  |                   |:  :|:
Db        760 ELVLLMERCWAQDPAERPDFGQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEER 819

Qy        440 HQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSP 499
               ||  |||:|   ||  | |  ||:|| :|: ||| | :||: |||||||||: :: :|
Db        820 TQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAESTP 879

Qy        500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL-HKESDTHAVQIALMALKMME 558
              :||:|:|| ||| ||       |||||||||||| |  ||  :    || :|| ||| ::
Db        880 MQVVTLLNDLYTCFDAIIDNFDVYKVETIGDAYMVVSGLPGRNGQRHAPEIARMALALL- 938

Qy        559 LSDEVMS------PHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
               | | |       || : :::|||:|:| | |||||:|||||||||: |  |:: ||
Db        939 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGQ 995

Qy        613 PRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              ||:|| ||   | :   |    |   |:
Db        996 ALKIHVSSTTKDALDELGCFQLELRGDVEM 1025
```

```
RESULT 14
ANPB_RAT
ID   ANPB_RAT       STANDARD;      PRT;  1047 AA.
AC   P16067;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B)
DE   (Guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide
DE   B-type receptor).
GN   NPR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89376566; PubMed=2570641;
RA   Schulz S., Singh S., Bellet R.A., Singh G., Tubb D.J., Chin H.,
RA   Garbers D.L.;
RT   "The primary structure of a plasma membrane guanylate cyclase
RT   demonstrates diversity within this new receptor family.";
RL   Cell 58:1155-1162(1989).
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF. SEEMS TO BE STIMULATED MORE
CC       EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M26896; AAA41205.1; -.
DR   PIR; A33300; OYRTBR.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001170; Ntpep_receptorN.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23   1047       ATRIAL NATRIURETIC PEPTIDE RECEPTOR B.
FT   DOMAIN      23    458       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   459    478       POTENTIAL.
FT   DOMAIN     479   1047       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     513    786       PROTEIN KINASE-LIKE.
FT   DOMAIN     861    991       GUANYLATE CYCLASE.
FT   DISULFID    75    101       BY SIMILARITY.
FT   DISULFID   439    439       INTERCHAIN (PROBABLE).
FT   DISULFID   448    448       INTERCHAIN (PROBABLE).
FT   CARBOHYD    24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1047 AA;  117126 MW;  5062C49228CC14A3 CRC64;

  Query Match           12.8%;  Score 460;  DB 1;  Length 1047;
  Best Local Similarity   41.9%;  Pred. No. 2.6e-24;
  Matches  113;  Conservative   39;  Mismatches   76;  Indels   42;  Gaps    6;

Qy        411 DVVLIGEQARAQD----------GLKKRLGK--------------------LKATLEQA 439
              ::||: |:  |||          |  :|  |                    |:  :|:
Db        760 ELVLLMERCWAQDPTERPDFGQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEER 819

Qy        440 HQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGFTAICSQCSP 499
               ||  |||:|   ||  | |  ||:|| :|: |||: | :||: |||||||||: :: :|
Db        820 TQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAESTP 879

Qy        500 LQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL-HKESDTHAVQIALMALKMME 558
              :||:|:|| ||| ||      |||||||||||| |  ||  :    || :|| ||| ::
Db        880 MQVVTLLNDLYTCFDAIIDNFDVYKVETIGDAYMVVSGLPGRNGQRHAPEIARMALALL- 938

Qy        559 LSDEVMS------PHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTLANKFESCSV 612
               | | |      || : :::|||:|:| | |||||:||||||||: |  |:: ||
Db        939 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGQ 995

Qy        613 PRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              ||:|| ||   | :   |   |   |:
Db        996 ALKIHVSSTTKDALDELGCFQLELRGDVEM 1025


RESULT 15
ANPA_RAT
ID   ANPA_RAT       STANDARD;      PRT;  1057 AA.
AC   P18910;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)
DE   (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide
DE   A-type receptor).
GN   NPR1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=89143770; PubMed=2563900;
RA   Chinkers M., Garbers D.L., Chang M.S., Lowe D.G., Chin H.,
RA   Goeddel D.V., Schulz S.;
RT   "A membrane form of guanylate cyclase is an atrial natriuretic
RT   peptide receptor.";
RL   Nature 338:78-83(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91056089; PubMed=1978722;
RA   Yamaguchi M., Rutledge L.J., Garbers D.L.;
RT   "The primary structure of the rat guanylyl cyclase A/atrial
RT   natriuretic peptide receptor gene.";
RL   J. Biol. Chem. 265:20414-20420(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91352095; PubMed=1679239;
RA   Duda T., Goraczniak R.M., Sharma R.K.;
RT   "Site-directed mutational analysis of a membrane guanylate cyclase
RT   cDNA reveals the atrial natriuretic factor signaling site.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 88:7882-7886(1991).
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X14773; CAA32881.1; -.
DR   EMBL; J05677; AAA41200.1; -.
DR   EMBL; M74535; AAA41202.1; -.
DR   PIR; S03348; OYRTR.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001170; Ntpep_receptorN.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL       1     28
FT   CHAIN       29   1057       ATRIAL NATRIURETIC PEPTIDE RECEPTOR A.
FT   DOMAIN      29    469       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   470    490       POTENTIAL.
FT   DOMAIN     491   1057       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     524    801       PROTEIN KINASE-LIKE.
FT   DOMAIN     872   1002       GUANYLATE CYCLASE.
FT   DISULFID    88    114       BY SIMILARITY.
FT   DISULFID   192    241       BY SIMILARITY.
FT   DISULFID   451    451       INTERCHAIN (PROBABLE).
FT   DISULFID   460    460       INTERCHAIN (PROBABLE).
FT   CARBOHYD    41     41       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   334    334       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   375    375       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   382    382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   366    366       Q -> H (IN REF. 3).
FT   CONFLICT   392    392       L -> P (IN REF. 3).
SQ   SEQUENCE   1057 AA;  118951 MW;  9EA9AE685AC05816 CRC64;

  Query Match             12.7%;  Score 457.5;  DB 1;  Length 1057;
  Best Local Similarity   44.9%;  Pred. No. 3.9e-24;
  Matches  102;  Conservative   40;  Mismatches   78;  Indels    7;  Gaps    3;

Qy        423 DGLKKRL----GKLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFS 478
              | |  |:      |:  :|:  ||  |||:|   ||  | |  ||:|| :|: |||: |
Db        810 DNLLSRMEQYANNLEELVEERTQAYLEEKRKAEALLYQILPHSVAEQLKRGETVQAEAFD 869

Qy        479 NVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQQCGELDVYKVETIGDAYCVAGGL 538
              :||: |||||||||: :: :|:||:|:|| ||| ||      ||||||||||| |  ||
Db        870 SVTIYFSDIVGFTALSAESTPMQVVTLLNDLYTCFDAVIDNFDVYKVETIGDAYMVVSGL 929

Qy        539 H-KESDTHAVQIALMALKMMEL--SDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYC 595
               :    || ::| ||| :::   |  :     | :::|||:|:| | |||||:||||||
Db        930 PVRNGQLHAREVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYC 989

Qy        596 LFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREEL 642
              |||: |  |:: ||     ||::|  |  :|:::  ||    |  |:
Db        990 LFGDTVNTASRMESNGEALKIHLSSETKAVLEEFDGFELELRGDVEM 1036
```

Search completed: June 27, 2003, 13:03:26
Job time : 13.0695 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 12:55:07 ; Search time 9.93048 Seconds
(without alignments)
2585.358 Million cell updates/sec

Title: US-09-762-767A-4
Perfect score: 3231
Sequence: 1 MYGFVNHALELLVIRNYGPE.....QVWFLSRKNTGTEETKQDDD 619

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3231	100.0	619	1	CYG1_HUMAN	Q02153 homo sapien
2	3206	99.2	619	1	CYG1_BOVIN	P16068 bos taurus
3	3188	98.7	619	1	CYG1_RAT	P20595 rattus norv
4	987	30.5	682	1	CYG2_RAT	P22717 rattus norv
5	927	28.7	617	1	CYG2_HUMAN	O75343 homo sapien
6	916	28.4	730	1	CYG4_RAT	Q9wvi4 rattus norv
7	909.5	28.1	732	1	CYG4_HUMAN	P33402 homo sapien
8	793	24.5	690	1	CYG3_RAT	P19686 rattus norv
9	789.5	24.4	690	1	CYG3_HUMAN	Q02108 homo sapien
10	770	23.8	691	1	CYG3_BOVIN	P19687 bos taurus
11	712	22.0	683	1	CYGH_DROME	Q07093 drosophila
12	481.5	14.9	1061	1	ANPA_HUMAN	P16066 homo sapien
13	468	14.5	1057	1	ANPA_RAT	P18910 rattus norv
14	466	14.4	1057	1	ANPA_MOUSE	P18293 mus musculu
15	454.5	14.1	1047	1	ANPB_RAT	P16067 rattus norv
16	453.5	14.0	1047	1	ANPB_HUMAN	P20594 homo sapien
17	450.5	13.9	1047	1	ANPB_BOVIN	P46197 bos taurus
18	438.5	13.6	1110	1	CYGD_BOVIN	P55203 bos taurus
19	433.5	13.4	1108	1	CYGE_MOUSE	P52785 mus musculu
20	433	13.4	1109	1	CYGD_CANFA	O19179 canis famil
21	428.5	13.3	1108	1	CYGE_RAT	P51840 rattus norv
22	425	13.2	1050	1	ANPB_ANGJA	P55202 anguilla ja
23	424	13.1	1108	1	CYGF_HUMAN	P51841 homo sapien
24	422.5	13.1	1103	1	CYGD_HUMAN	Q02846 homo sapien
25	420	13.0	1108	1	CYGF_RAT	P51842 rattus norv
26	419	13.0	1103	1	CYGF_BOVIN	O02740 bos taurus
27	415.5	12.9	1110	1	CYGX_RAT	P51839 rattus norv
28	410	12.7	1137	1	CYG1_CAEEL	Q09435 caenorhabdi
29	409	12.7	433	1	KSGC_RAT	P55205 rattus norv
30	405.5	12.6	1073	1	HSER_HUMAN	P25092 homo sapien
31	405	12.5	1073	1	HSER_PIG	P55204 sus scrofa
32	404	12.5	1072	1	HSER_RAT	P23897 rattus norv
33	391	12.1	1125	1	CYGS_STRPU	P16065 strongyloce
34	389	12.0	1076	1	HSER_CAVPO	P70106 cavia porce
35	345.5	10.7	443	1	CYA1_MYCTU	O30820 mycobacteri
36	309.5	9.6	1134	1	CYA1_BOVIN	P19754 bos taurus
37	308.5	9.5	1248	1	CYA8_RAT	P40146 rattus norv
38	302.5	9.4	1251	1	CYA8_HUMAN	P40145 homo sapien
39	302.5	9.4	1305	1	CYA9_XENLA	P98999 xenopus lae
40	301.5	9.3	839	1	CYA1_HUMAN	Q08828 homo sapien
41	301.5	9.3	1249	1	CYA8_MOUSE	P97490 mus musculu
42	300.5	9.3	2248	1	CYA1_DROME	P32870 drosophila
43	299	9.3	858	1	CYAG_DICDI	Q03101 dictyosteli
44	293	9.1	1144	1	CYA3_HUMAN	O60266 homo sapien
45	292.5	9.1	1334	1	CYA9_CHICK	Q9dgg6 gallus gall

ALIGNMENTS

RESULT 1
CYG1_HUMAN
ID CYG1_HUMAN STANDARD; PRT; 619 AA.
AC Q02153;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
DE (Soluble guanylate cyclase small subunit) (GCS-beta-3).
GN GUCY1B1 OR GUCY1B3 OR GUC1B3 OR GUCSB3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92316204; PubMed=1352257;
RA Giuili G., Scholl U., Bulle F., Guellaeen G.;
RT "Molecular cloning of the cDNAs coding for the two subunits of
RT soluble guanylyl cyclase from human brain.";
RL FEBS Lett. 304:83-88(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM HSGC-2).
RC TISSUE=Kidney;
RA Gansemans Y., Brouckaert P., Fiers W.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 337-545, AND ALTERNATIVE SPLICING.
RC TISSUE=Lung;
RX MEDLINE=92008652; PubMed=1680753;
RA Chhajilani V., Fraendberg P.-A., Ahlner J., Axelsson K.L.,
RA Wikberg J.E.S.;
RT "Heterogeneity in human soluble guanylate cyclase due to alternative
RT splicing.";
RL FEBS Lett. 290:157-158(1991).
CC -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HSGC-1 (SHOWN HERE) AND HSGC-2;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.

```
DR   EMBL; X66533; CAA47144.1; -.
DR   EMBL; AF020340; AAB94877.1; -.
DR   PIR; S23097; S23097.
DR   HSSP; P16068; 1AWN.
DR   Genew; HGNC:4687; GUCY1B3.
DR   MIM; 139397; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Alternative splicing.
FT   DOMAIN      421    554       GUANYLATE CYCLASE.
FT   VARSPLIC    393    425       MISSING (IN ISOFORM HSGC-2).
SQ   SEQUENCE   619 AA;  70514 MW;  231E4E660DE02AA1 CRC64;

  Query Match             100.0%;  Score 3231;  DB 1;  Length 619;
  Best Local Similarity   100.0%;  Pred. No. 2e-215;
  Matches  619;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              ||||||||||||||||||
Db        601 VWFLSRKNTGTEETKQDDD 619


RESULT 2
CYG1_BOVIN
ID   CYG1_BOVIN     STANDARD;      PRT;   619 AA.
AC   P16068;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
DE   (Soluble guanylate cyclase small subunit).
GN   GUCY1B1 OR GUCY1B3 OR GUC1B3.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=89031214; PubMed=2903071;
RA   Koesling D., Herz J., Gausepohl H., Niroomand F., Hinsch K.-D.,
RA   Muelsch A., Boehme E., Schultz G., Frank R.;
RT   "The primary structure of the 70 kDa subunit of bovine soluble
RT   guanylate cyclase.";
RL   FEBS Lett. 239:29-34(1988).
RN   [2]
RP   3D-STRUCTURE MODELING OF 412-572.
RX   MEDLINE=98054247; PubMed=9391039;
RA   Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.;
RT   "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
RT   and mutational analysis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: LUNG AND BRAIN.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00770; CAA68739.1; -.
DR   PIR; S01653; OYBO70.
DR   PDB; 1AWN; 28-JAN-98.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; 3D-structure.
FT   DOMAIN      421    554       GUANYLATE CYCLASE.
SQ   SEQUENCE   619 AA;  70502 MW;  8EFB14952B80F344 CRC64;

  Query Match             99.2%;  Score 3206;  DB 1;  Length 619;
  Best Local Similarity   99.0%;  Pred. No. 1e-213;
  Matches  613;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCTDADKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
```

```
Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKSECEDELTGTE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMTPENSDPQFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              |||||||||||||||:||::
Db        601 VWFLSRKNTGTEETEQDEN 619


RESULT 3
CYG1_RAT
ID   CYG1_RAT       STANDARD;      PRT;   619 AA.
AC   P20595;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
DE   (Soluble guanylate cyclase small subunit).
GN   GUCY1B1 OR GUCY1B3 OR GUC1B3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89087429; PubMed=2905128;
RA   Nakane M., Saheki S., Kuno T., Ishii K., Murad F.;
RT   "Molecular cloning of a cDNA coding for 70 kilodalton subunit of
RT   soluble guanylate cyclase from rat lung.";
RL   Biochem. Biophys. Res. Commun. 157:1139-1147(1988).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: LUNG AND BRAIN.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22562; AAA41204.1; -.
DR   PIR; A31871; OYRTB1.
DR   HSSP; P16068; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis.
FT   DOMAIN      421    554       GUANYLATE CYCLASE.
SQ   SEQUENCE   619 AA;  70477 MW;  BCC9850DE4504226 CRC64;

  Query Match             98.7%;  Score 3188;  DB 1;  Length 619;
  Best Local Similarity   98.5%;  Pred. No. 1.8e-212;
  Matches  610;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              |||||||||||||||||||||||||||||||:|||||||||| ||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDQEGQFLVRIIYDHSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRSEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQDSRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        241 PQLQPGKCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGAE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              |||||||||||||| ||::
Db        601 VWFLSRKNTGTEETNQDEN 619


RESULT 4
CYG2_RAT
ID   CYG2_RAT       STANDARD;      PRT;   682 AA.
AC   P22717;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-2 chain (EC 4.6.1.2) (GCS-beta-2).
GN   GUCY1B2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Kidney;
RX   MEDLINE=91105012; PubMed=1980215;
RA   Yuen P.S.T., Potter L.R., Garbers D.L.;
RT   "A new form of guanylyl cyclase is preferentially expressed in rat
RT   kidney.";
RL   Biochemistry 29:10872-10878(1990).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: KIDNEY AND LIVER.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57507; AAA41207.1; -.
DR   PIR; A36228; OYRTB2.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis.
FT   DOMAIN      408    536       GUANYLATE CYCLASE.
SQ   SEQUENCE   682 AA;  76196 MW;  98C173C1A1CC7715 CRC64;

  Query Match             30.5%;  Score 987;  DB 1;  Length 682;
  Best Local Similarity   36.9%;  Pred. No. 1.4e-60;
  Matches  225;  Conservative  104;  Mismatches  195;  Indels   86;  Gaps   12;

Qy         66 ILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDA 125
              ||::||: || ||: |||| :|| || |: ||::|||||| :||  |  | |||||  :
Db          4 ILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDALHSYLALSYQEMNAPSFRVEEG 63

Qy        126 EKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC------DHTQF 179
                |  ::|||||:| ||  || |||: ||:    |:: | ::   |||       :|  |
Db         64 ADG-AMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMSILDM-NEEVERTGKKEHVVF 121

Qy        180 LIEEKESKE------------EDFYED------------------------------- 194
              |: :|  ::             ||   |
Db        122 LVVQKAHRQIRGAKASRPQGSEDSQADQEALQGTLLRMKERYLNIPVCPGEKSHSTAVRA 181

Qy        195 ---------LDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQP 245
                      | |:    :   :    || ||||||:||  | | :| |  | : :| :
Db        182 SVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFHIVFDEALRVKQAGVNIQKYVPGILT 241

Qy        246 GNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLR 305
               :|   ||:: | :  :   |   ||: |||:::: ::   :           |:
Db        242 QKFALDEYFSIIHPQVTFNISSICKFINSQFVLKTRKEMMPKAR---------KSQPMLK 292

Qy        306 LKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKL 365
              |:||||::    ::|:|||:| :| :|    ::|||  || ||||:|| :|   | :|
Db        293 LRGQMIWMESLRCMIFMCSPNVRSLQELEESKMHLSDIAPHDTTRDLILLNQQRLAEMEL 352

Qy        366 TQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFS 425
              : :||   : |::    |  ||||:||||::||  |||:|:  | | |  ::  |||||
Db        353 SCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVANQLKEGRKVAAGEFETCTILFS 412

Qy        426 GIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGL 485
               :| |   |   |: |  ::|||:|| :|::|| ||       ||||||:|| || | |:
Db        413 DVVTFTNIC---AACE-PIQIVNMLNSMYSKFDRLTSVHD---VYKVETIGDAYMVVGGV 465

Qy        486 PEPCIHHARSICHLALDMMEIAGQVQ--VDGESVQITIGIHTGEVVTGVIGQRMPRYCLF 543
                | |   ||: : : || |   | :|   | || :|| :||||| |: ||:| :|||||||
Db        466 PVPVESHAQRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVVGDKMPRYCLF 525

Qy        544 GNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWF 603
              |:|||  || |: |   |:::|   :| |       :  | :  || : :||| : |  :|
Db        526 GDTVNTASRMESHGLPSKVHLSPTAHRAL-----KNKGFEIVRRGEIEVKGKGK-MTTYF 579

Qy        604 LSRKNTGTEE 613
              | :   ||:
Db        580 LIQNLNATED 589


RESULT 5
CYG2_HUMAN
ID   CYG2_HUMAN     STANDARD;      PRT;   617 AA.
AC   O75343; Q9NZ64;
DT   15-JUL-1999 (Rel. 38, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, beta-2 chain (EC 4.6.1.2) (GCS-beta-2).
GN   GUCY1B2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND REVISION TO 175.
RC   TISSUE=Kidney;
RX   MEDLINE=20241821; PubMed=10777682;
RA   Behrends S., Vehse K.;
RT   "The beta(2) subunit of soluble guanylyl cyclase contains a
RT   human-specific frameshift and is expressed in gastric carcinoma.";
RL   Biochem. Biophys. Res. Commun. 271:64-69(2000).
RN   [2]
RP   SEQUENCE OF 95-257 FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=99107820; PubMed=9889008;
RA   Behrends S., Kazmierczak B., Steenpa A., Knauf B., Bullerdiek J.,
RA   Scholz H., Eiberg H.;
RT   "Assignment of GUCY1B2, the gene coding for the beta2 subunit of human
RT   guanylyl cyclase to chromosomal band 13q14.3 between markers D13S168
RT   and D13S155.";
RL   Genomics 55:126-127(1999).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF038499; AAD09440.2; -.
DR   EMBL; AF218383; AAF66105.1; -.
DR   HSSP; P19687; 1AWN.
DR   Genew; HGNC:4686; GUCY1B2.
DR   MIM; 603695; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
```

```
KW   Lyase; cGMP synthesis.
FT   DOMAIN      391    519       GUANYLATE CYCLASE.
SQ   SEQUENCE   617 AA;  70367 MW;  1B6C965CAB9A5E3E CRC64;

  Query Match             28.7%;  Score 927;  DB 1;  Length 617;
  Best Local Similarity   36.6%;  Pred. No. 1.6e-56;
  Matches  218;  Conservative   96;  Mismatches  195;  Indels   86;  Gaps   11;

Qy         81 SGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSERE 140
              |||| :|| || |: ||::|||||| :||  |  | |||||      || : |||||:|
Db          2 SGYDRMLRTLGGNLMEFIENLDALHSYLALSYQEMNAPSFRVERGADGK-MFLHYYSDRS 60

Qy        141 GLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC------DHTQFLIEEKES-------- 186
              ||  || |||: ||:     :: | ::   |||       :|  ||| :|
Db         61 GLCHIVPGIIEAVAKDFFDIDVIMDILDM-NEEVERTGKKEHVVFLIVQKAHRKMRKTKP 119

Qy        187 --------------------KEEDFYED----------------------LDRFEE 200
                                 | :: | :                      :: ||
Db        120 KRLQDSQGMERDQEALQAAFLKMKEKYLNVSACPVKKSHWDVVRSIVMFGKGHLMNTFEP 179

Qy        201 NGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPH 260
                :   |   ||| ||||||:||  | | |     | : :| ||  |  |   ||:: |
Db        180 IYPERLWIEEKTFCNAFPFHIVFDESLQVKQARVNIQKYVPGLQTQNIQLDEYFSIIHPQ 239

Qy        261 IDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQMIYLPEADSIL 320
              :  :   |   ||: |||:::  :: |             : |:|:||||::     ::
Db        240 VTFNIFSIRRFINSQFVLKTRREMMPV---------AWQSRTTLKLQGQMIWMESMWCMV 290

Qy        321 FLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTL 380
              :|||| : :| :|    ::|||| :| ||||:|| :|   | :|: :||   : ||:
Db        291 YLCSPKLRSLQELEELNMHLSDIAPNDTTRDLILLNQQRLAEIELSNQLERKKEELQVLS 350

Qy        381 RALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASG 440
              : |  |||||:||||::||  |||:||  : | |  : : ||||| :| |   |:
Db        351 KHLAIEKKKTETLLYAMLPKHVANQLREGKKVAAGEFKSCTILFSDVVTFTNICTACE-- 408

Qy        441 EGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLA 500
              ::|||:|| :|::|| ||       ||||||:|| || | |:| |  :||: : : |
Db        409 --PIQIVNVLNSMYSKFDRLTSVHA---VYKVETIGDAYMVVGGVPVPIGNHAQRVANFA 463

Qy        501 LDMMEIAGQV--QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGE 558
              | |   | :|   | || :|: :||||| |:  |:| :||||||||:|||  || |: |
Db        464 LGMRISAKEVTNPVTGEPIQLRVGIHTGPVLADVVGDKMPRYCLFGDTVNTASRMESHGL 523

Qy        559 KGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSRKNTGTEE 613
                |:::|   || |       :  | :  || : :|||   |  :|| :    ||:
Db        524 PNKVHLSPTAYRAL-----KNQGFKIIERGEIEVKGKGR-MTTYFLIQNLNATED 572


RESULT 6
CYG4_RAT
ID   CYG4_RAT       STANDARD;      PRT;   730 AA.
AC   Q9WVI4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-2 chain (EC 4.6.1.2) (GCS-alpha-2).
GN   GUCY1A2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar Kyoto; TISSUE=Aorta;
RX   MEDLINE=20571097; PubMed=11121588;
RA   Koglin M., Behrends S.;
RT   "Cloning and functional expression of the rat alpha(2) subunit of
RT   soluble guanylyl cyclase.";
RL   Biochim. Biophys. Acta 1494:286-289(2000).
CC   -!- FUNCTION: HAS GUANYLYL CYCLASE ON BINDING TO THE BETA-1 SUBUNIT.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF109963; AAD42949.2; -.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family.
FT   DOMAIN      519    646       GUANYLATE CYCLASE.
FT   DOMAIN       50     72       ALA-RICH.
FT   DOMAIN       64     72       POLY-ALA.
SQ   SEQUENCE   730 AA;  81786 MW;  B8D790BFF81FB8F9 CRC64;

  Query Match             28.4%;  Score 916;  DB 1;  Length 730;
  Best Local Similarity   38.0%;  Pred. No. 1.2e-55;
  Matches  210;  Conservative  116;  Mismatches  167;  Indels   60;  Gaps   16;

Qy         51 LVAAASKVLNLNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLAT 110
              ::   : || |   || : ||: ||  | :   : :|| :|| :::|    ||| :|: |
Db        152 ILRCTANVLGLKFQEIQERFGEEFFKICFDEN-ERVLRAVGSTLQDFFNGFDALLEHIRT 210

Qy        111 IY---PGMRAPSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVI 167
               :     : :||| | :  :|  | |||:     :   ::|:||   ::|:   :::: |
Db        211 SFGKQATLESPSFLCKELPEGT-LKLHYFHPHHTVGFAMLGMIKAAGKRIYHLNVEVEQI 269

Qy        168 QQR----------NEECDHTQFLIEEKESKEEDFYEDLDRFEENGTQ----ESRISPYTF 213
              :              | |    |||:| |: :      : :    ||      : |||  ||
Db        270 ENEKFCSDGSTPSNYSC--LTFLIKECETTQ------ITKNIPQGTSQIPTDLRISINTF 321

Qy        214 CKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQPGNCSLL---SVFSLVRPHIDISFHGILS 270
              |: ||||::|| ::|| | |  : :   ||:  |  :|     | :| | :: :|  :|
Db        322 CRTFPFHLMFDPNMVVLQLGEGLRK---QLRCDNHKVLKFEDCFEIVSPKVNATFDRVLL 378

Qy        271 HINTVFVLRSKEGLLDVEKLECEDELTGT--EISCLRLKGQMIYLPEADSILFLCSPSVM 328
              ::| ||:|:|              | :||  |   : :||||::||:::|||| || |
Db        379 RLSTPFVIRTK------------PEASGTDNEDKVMEIKGQMIHVPESNAILFLGSPCVD 426

Qy        329 NLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKK 388
              ||:|  |||:|||||:||||||::|:||| : :  | : :: |   |: | :|||:|||
Db        427 KLDELIGRGLHLSDIPIHDATRDVILVGEQAKAQDGLKKRMDKLKATLEKTHQALEEEKK 486

Qy        389 KTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVN 448
              ||  |||| | || :|  :: | |:::|:||:||| |||| | |::       |::::
Db        487 KTVDLLYSIFPGDVAQQLWQRQQVQARKFDDVTMLFSDIVGFTAICAQCT----PMQVIS 542

Qy        449 LLNDLYTRFDTLTDSRKNPF--VYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEI 506
              :||:||||||       :  |  :|||||:|| |   |||    : ||: |  :|| |||:
Db        543 MLNELYTRFD-----HQCGFLDIYKVETIGDAYCVASGLHRKSLCHAKPIALMALKMMEL 597

Qy        507 AGQVQV-DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVS 565
              : :|   ||  :|: ||||:| |: ||:| ||||||||| | | |: |:     :||:|
Db        598 SEEVLTPDGRPIQMRIGIHSGSVLAGVVGVRMPRYCLFGNNVTLASKFESGSHPRRINIS 657

Qy        566 EYTYRCLMSPENS 578
                ||: |:  |:|
Db        658 PTTYQ-LLKREDS 669
```

```
RESULT 7
CYG4_HUMAN
ID   CYG4_HUMAN     STANDARD;      PRT;   732 AA.
AC   P33402;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-2 chain (EC 4.6.1.2) (GCS-alpha-2).
GN   GUCY1A2 OR GUC1A2 OR GUCSA2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92070494; PubMed=1683630;
RA   Harteneck C., Wedel B., Koesling D., Malkewitz J., Boehme E.,
RA   Schultz G.;
RT   "Molecular cloning and expression of a new alpha-subunit of soluble
RT   guanylyl cyclase. Interchangeability of the alpha-subunits of the
RT   enzyme.";
RL   FEBS Lett. 292:217-222(1991).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA 2-I).
RX   MEDLINE=95403342; PubMed=7673142;
RA   Behrends S., Harteneck C., Schultz G., Koesling D.;
RT   "A variant of the alpha 2 subunit of soluble guanylyl cyclase
RT   contains an insert homologous to a region within adenylyl cyclases
RT   and functions as a dominant negative protein.";
RL   J. Biol. Chem. 270:21109-21113(1995).
CC   -!- FUNCTION: HAS GUANYLYL CYCLASE ON BINDING TO THE BETA-1 SUBUNIT.
CC       THE ALTERNATIVE SPLICED ISOFORM, ALPHA-2-I, ACTS AS A NEGATIVE
CC       REGULATOR OF GUANYLYL CYLASE ACTIVITY AS IT FORMS NON-FUNCTIONAL
CC       HETERODIMERS WITH THE BETA SUBUNITS.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-2 (SHOWN HERE) AND
CC       ALPHA-2-I; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALPHA-2 ISOFORM IS EXPRESSED IN FETAL BRAIN,
CC       LIVER, COLON, ENDOTHELIUM AND TESTIS, WHILE THE ALPHA-2-I FORM
CC       IS EXPRESSED ONLY IN LIVER, COLON AND ENDOTHELIUM.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63282; CAA44921.1; -.
DR   EMBL; Z50053; CAA90393.1; -.
DR   PIR; S18325; S18325.
DR   HSSP; P19687; 1AWN.
DR   Genew; HGNC:4684; GUCY1A2.
DR   MIM; 601244; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family; Alternative splicing.
FT   DOMAIN      521     648       GUANYLATE CYCLASE.
FT   DOMAIN       51      76       ALA-RICH.
FT   DOMAIN       51      58       POLY-ALA.
FT   VARSPLIC    612     612       Q -> QPQRSELLFSFPVSIQLVPDQHQSETDLGTEK
FT                                 (IN ISOFORM ALPHA 2-I).
SQ   SEQUENCE   732 AA;  81749 MW;  79465A7D3FE52DB7 CRC64;

  Query Match             28.1%;  Score 909.5;  DB 1;  Length 732;
  Best Local Similarity   37.7%;  Pred. No. 3.3e-55;
  Matches  208;  Conservative  111;  Mismatches  175;  Indels   57;  Gaps   14;

Qy         51 LVAAASKVLNLNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLAT 110
              ::   : :| |   || : ||: ||  |     : :|| :|  :::|   ||| :|: |
Db        154 ILQCTANILGLKFEEIQKRFGEEFFNICFHEN-ERVLRAVGGTLQDFFNGFDALLEHIRT 212

Qy        111 IY---PGMRAPSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVI 167
               :     : :||| | :  :|  |:|||:     :   ::|:|| ::|:  :|::|
Db        213 SFGKQATLESPSFLCKELPEGT-LMLHYFHPHHIVGFAMLGMIKAAGKKIY--RLDVEVE 269

Qy        168 QQRNEE----------CDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAF 217
              |  ||:          |    |||  || :  :  ::| :       : |||  |||:||
Db        270 QVANEKLCSDVSNPGNCSCLTFLI--KECENTNIMKNLPQGTSQVPADLRISINTFCRAF 327

Qy        218 PFHIIFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSV------FSLVRPHIDISFHGILSH 271
              |||::||  : | | |  : : |       |    |      | :| | :: :|  :|
Db        328 PFHLMFDPSMSVLQLGEGLRKQL------RCDTHKVLKFEDCFEIVSPKVNATFERVLLR 381

Qy        272 INTVFVLRSKEGLLDVEKLECEDELTGTE--ISCLRLKGQMIYLPEADSILFLCSPSVMN 329
              ::| ||:|:|           | :|:|     : :|||||::||::||||| || |
Db        382 LSTPFVIRTK-----------PEASGSENKDKVMEVKGQMIHVPESNSILFLGSPCVDK 429

Qy        330 LDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKK 389
              ||:|  |||:|||||:||||||::|:||| : :  | : :: |   |: | :|||:||||
Db        430 LDELMGRGLHLSDIPIHDATRDVILVGEQAKAQDGLKKRMDKLKATLERTHQALEEEKKK 489

Qy        390 TDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNL 449
              |  ||||: |  || :|   : | |:::|:||:||| |||| | |::       |:::::
Db        490 TVDLLYSIFPGDVAQQLWQGQQVQARKFDDVTMLFSDIVGFTAICAQCT----PMQVISM 545

Qy        450 LNDLYTRFDTLTDSRKNPF--VYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIA 507
              ||:||||||       :  |  :|||||:|| |   :||    : ||: |  :|| |||::
Db        546 LNELYTRFD-----HQCGFLDIYKVETIGDAYCVAAGLHRKSLCHAKPIALMALKMMELS 600

Qy        508 GQVQV-DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSE 566
               :|   ||  :|: ||||:| |: ||:| ||||||||| | | |: |:     :||||
Db        601 EEVLTPDGRPIQMRIGIHSGSVLAGVVGVRMPRYCLFGNNVTLASKFESGSHPRRINVSP 660

Qy        567 YTYRCLMSPEN 577
              ||: |   |:
Db        661 TTYQLLKREES 671


RESULT 8
CYG3_RAT
ID   CYG3_RAT       STANDARD;      PRT;   690 AA.
AC   P19686;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)
DE   (Soluble guanylate cyclase large subunit).
GN   GUCY1A1 OR GUC1A1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=91009100; PubMed=1698769;
RA   Nakane M., Arai K., Saheki S., Kuno T., Buechler W., Murad F.;
RT   "Molecular cloning and expression of cDNAs coding for soluble
RT   guanylate cyclase from rat lung.";
RL   J. Biol. Chem. 265:16841-16845(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=Sprague-Dawley; TISSUE=Striatum;
RX   MEDLINE=97151525; PubMed=8997507;
RA   Smigrodzki R.M., Levitt P.;
RT   "The alpha 1 subunit of soluble guanylyl cyclase is expressed
RT   prenatally in the rat brain.";
RL   Brain Res. Dev. Brain Res. 97:226-234(1996).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57405; AAA41206.1; -.
DR   EMBL; U60835; AAB17953.1; -.
DR   PIR; A38297; OYRTA1.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family.
FT   DOMAIN      480    607       GUANYLATE CYCLASE.
SQ   SEQUENCE   690 AA;  77566 MW;  E4819B2CA4F86401 CRC64;

  Query Match             24.5%;  Score 793;  DB 1;  Length 690;
  Best Local Similarity   34.7%;  Pred. No. 3.3e-47;
  Matches  214;  Conservative  100;  Mismatches  215;  Indels   88;  Gaps   20;

Qy         27 KEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEILQ-MFGKMFFVFCQESGYDT 85
              ||  :   | : | ||| :::     :||   |      |:|:   |:  |  | |   :
Db        102 KENRNSSEKEDLERIIAEEA-----IAAGVPV------EVLKDSLGEELFKICYEED-EH 149

Qy         86 ILRVLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSERE 140
              || |:|  :::|| :   |     |   |     :   |  | | ::   | ::|:  :
Db        150 ILGVVGGTLKDFLNSFSTLLKQSSHCQEAERRGRLEDASILCLDKDQ-DFLNVYYFFPKR 208

Qy        141 GLQDIVIGIIKTVAQQIHGTEIDMKVIQQ-RNEECDHTQFL--------IEEKESKEEDF 191
                 :: ||||  |: :: : ::: ::      ||  |:|:        :  | :|
Db        209 TTALLLPGIIKAAARILYESHVEVSLMPPCFRSEC--TEFVNQPYLLYSVHVKSTKPS-- 264

Qy        192 YEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCS 249
                       | | : | : ||| |||| : |||| : | || | |:: :    |  :
Db        265 ------LSPGKPQSSLVIPTSLFCKTFPFHFMLDRDLAILQLGNGIRRLVNKRDFQGKPN 318

Qy        250 LLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQ 309
                | :: | |: :| ||:: :|  ||:| :             | |       : ||||
Db        319 FEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRW----------DNLVKKSSRVMDLKGQ 368

Qy        310 MIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQEL 369
              |||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        369 MIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRL 428

Qy        370 EILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVG 429
                 |   |:   :|||:|||||  || |: |  || :|   : | ||::: ||:||| |||
Db        429 GKLKATLEHAHQALEEEKKKTVDLLCSIFPSEVAQQLWQGQIVQAKKFNEVTMLFSDIVG 488

Qy        430 FNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPC 489
              | | ||: :     :::: :|| ||||||          ||||||:|| |    ||
Db        489 FTAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHRES 541

Qy        490 IHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVN 548
                ||  |  :|| |||:: :|    || ::: ||:|:|  |  ||:| :|||||||| |
Db        542 DTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVT 601

Qy        549 LTSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQF--------HL----EH 586
              | :: |:     ||||  ||| |           | |   | |        |     :|
Db        602 LANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSDIPGICHFLDAYQH 661

Qy        587 RGPVSMKGKKEPMQVWF 603
              :|| |         : ||
Db        662 QGPNS--------KPWF 670


RESULT 9
CYG3_HUMAN
ID   CYG3_HUMAN     STANDARD;      PRT;   690 AA.
AC   Q02108; O43843;
DT   01-JUL-1993 (Rel. 26, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)
DE   (Soluble guanylate cyclase large subunit) (GCS-alpha-3).
GN   GUCY1A1 OR GUCY1A3 OR GUC1A3 OR GUCSA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=92316204; PubMed=1352257;
RA   Giuili G., Scholl U., Bulle F., Guellaeen G.;
RT   "Molecular cloning of the cDNAs coding for the two subunits of
RT   soluble guanylyl cyclase from human brain.";
RL   FEBS Lett. 304:83-88(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Gansemans Y., Brouckaert P., Fiers W.;
RT   "Human soluble guanylate cyclase large subunit mRNA, alpha3-like.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98416113; PubMed=9742212;
RA   Zabel U., Weeger M., La M., Schmidt H.H.;
RT   "Human soluble guanylate cyclase: functional expression and revised
RT   isoenzyme family.";
RL   Biochem. J. 335:51-57(1998).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66534; CAA47145.1; -.
DR   EMBL; U58855; AAB94794.1; -.
DR   EMBL; Y15723; CAA75738.1; -.
DR   PIR; S23098; S23098.
DR   HSSP; P19687; 1AWN.
DR   Genew; HGNC:4685; GUCY1A3.
```

```
DR   MIM; 139396; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family.
FT   DOMAIN      481    608       GUANYLATE CYCLASE.
FT   CONFLICT    124    127       AAGV -> QQS (IN REF. 1).
FT   CONFLICT    131    184       VIKESLGEEVFKICYEEDENILGVVGGTLKDFLNSFSTLLK
FT                                QSSHCQEAGKRGR -> LSKNLLVKRFLKYVTRKMKTSLGW
FT                                LEAPLKIFKQLQYPSETEQPLPRSRKKGQ (IN REF.
FT                                1).
FT   CONFLICT    322    322       MISSING (IN REF. 1).
FT   CONFLICT    529    535       GDAYCVA -> AMPIVWL (IN REF. 1).
FT   CONFLICT    678    690       GNANFLGKASGID -> ASQFFRQSIRNRLATYIPIYKSLG
FT                                FDSLKMCRASESTLGIVDG (IN REF. 1).
SQ   SEQUENCE   690 AA;  77452 MW;  DA1E14A5E11451CF CRC64;

  Query Match             24.4%;  Score 789.5;  DB 1;  Length 690;
  Best Local Similarity   35.1%;  Pred. No. 5.8e-47;
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88
              :: |  : | |  :: :     |||    |       |:: :  |:  |  | |   : ||
Db        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153

Qy         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143
               |:|  :::|| :   |      |   |      :   |  | | |    | ::|:  :
Db        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212

Qy        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ-----------FLIEEKESKEEDFY 192
               |: ||||  |  :: ||::: ::      | |              : :  | :|
Db        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265

Qy        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250
                    :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320

Qy        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310
                 | :: | |: :| ||:: :|  ||:| :     |:|               : |||||
Db        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370

Qy        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370
               ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430

Qy        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430
               |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490

Qy        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490
               | ||: :     :::: :|| ||||||          ||||||:|| |    || :
Db        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543

Qy        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549
               ||  |  :|| |||:: :|    || ::: ||:|:| |  ||:| :|||||||| | |
Db        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603

Qy        550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588
               :: |:     |||||  ||| |           | |   | |  | |
Db        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652


RESULT 10
CYG3_BOVIN
ID   CYG3_BOVIN     STANDARD;      PRT;   691 AA.
AC   P19687;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)
DE   (Soluble guanylate cyclase large subunit).
GN   GUCY1A1 OR GUC1A1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Adrenal medulla;
RX   MEDLINE=90306336; PubMed=1973124;
RA   Koesling D., Harteneck C., Humbert P., Bosserhoff A., Frank R.,
RA   Schultz G., Boehme E.;
RT   "The primary structure of the larger subunit of soluble guanylyl
RT   cyclase from bovine lung. Homology between the two subunits of the
RT   enzyme.";
RL   FEBS Lett. 266:128-132(1990).
RN   [2]
RP   3D-STRUCTURE MODELING OF 472-628.
RX   MEDLINE=98054247; PubMed=9391039;
RA   Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.;
RT   "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
RT   and mutational analysis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC       MAGNESIUM OR MANGANESE IONS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC       FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54014; CAA37960.1; -.
DR   PIR; S10713; OYBO77.
DR   PDB; 1AWN; 28-JAN-98.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family; 3D-structure.
FT   DOMAIN      482    609       GUANYLATE CYCLASE.
SQ   SEQUENCE   691 AA;  77532 MW;  5D1FE4D2204E8683 CRC64;

  Query Match             23.8%;  Score 770;  DB 1;  Length 691;
  Best Local Similarity   33.9%;  Pred. No. 1.3e-45;
  Matches  205;  Conservative   90;  Mismatches  200;  Indels  110;  Gaps   16;

Qy         27 KEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDT 85
               || :   | :   :|: | :     :||   |       ||: :  |:  |  | | :
Db        104 KENRKSLEREDFEKIVVDQA-----IAAGVPV------EIIKESLGEELFKICYEED-EY 151

Qy         86 ILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDAEKGKG--------------- 130
               || |:|  :::|| :   |              |  | :||| ||
Db        152 ILGVVGGTLKDFLNSFSTL----------LKQSSHCQEAEK-KGRFEDASILCLDKDPD 199

Qy        131 -LILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFLIEEKESKEE 189
                | ::|:  :     |: |||| |: :: ||:::     |
Db        200 VLYVYYFFPKRITSLILPGIIKAAARILYETEVEVSSTPSR------------------ 240

Qy        190 DFYEDLDRFEENGTQ------------------ESRISPYT-FCKAFPFHIIFDRDLVV 229
               |::|   | :   :                  | : | : ||| |||| : |||: :
Db        241 -FHQDCREFVDQPCELYSVHIRSARPHPPPGKPVSSLVIPASLFCKTFPFHFMLDRDMSI 299
```

```
Qy        230 TQCGNAIYRVLPQLQ-PGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVE 288
Db        300 LQLGHGIRRLMSRRDVQGKPHFDEYFEILTPKISQTFSGIMTMLNMQFLVRVRRW----- 354

Qy        289 KLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDA 348
Db        355 -----DNSMKKSSRVMDLKGQMIYMVESSSILFLGSPCVDRLEDFTGRGLYLSDIPIHNA 409

Qy        349 TRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRH 408
Db        410 LRDVVLIGEQARAQDGLKKRLGKLKATLEQAHQALEEEKRKTVDLLCSIFPSEVARQLWQ 469

Qy        409 KRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPF 468
Db        470 GHAVQAKRFGNVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFDRQCGELD--- 522

Qy        469 VYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGE 527
Db        523 VYKVETIGDAYCVAGGLHKESDTHAVQIALMALKMMELSHEVVSPHGEPIKMRIGLHSGS 582

Qy        528 VVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL----------MSPEN 577
Db        583 VFAGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREE 642

Qy        578 SDPQF 582
Db        643 LPPNF 647
```

```
RESULT 11
CYGH_DROME
ID   CYGH_DROME     STANDARD;      PRT;   683 AA.
AC   Q07093;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Head-specific guanylate cyclase (EC 4.6.1.2).
GN   GYC-ALPHA-99B OR GYC-ALPHA-63A OR GYC OR DGC1.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Canton-S; TISSUE=Head;
RX   MEDLINE=93203896; PubMed=8095978;
RA   Yoshikawa S., Miyamoto I., Aruga J., Furuichi T., Okano H.,
RA   Mikoshiba K.;
RT   "Isolation of a Drosophila gene encoding a head-specific guanylyl
RT   cyclase.";
RL   J. Neurochem. 60:1570-1573(1993).
CC   -!- FUNCTION: MAY HAVE A ROLE IN PHOTOTRANSDUCTION. A SECOND SUBUNIT
CC       MAY BE REQUIRED FOR ENZYME ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBUNIT: DIMER (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- TISSUE SPECIFICITY: HEAD, WHERE IT IS PREFERENTIALLY EXPRESSED IN
CC       THE CNS AND THE RETINA. NOT FOUND IN BODIES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S57126; AAB25820.1; -.
DR   PIR; JH0810; JH0810.
DR   HSSP; P19687; 1AWN.
DR   FlyBase; FBgn0013972; Gyc-alpha-99B.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cGMP synthesis; Multigene family; Vision.
FT   DOMAIN      465    591       GUANYLATE CYCLASE.
SQ   SEQUENCE   683 AA;  75906 MW;  12D67C4B79A07C8D CRC64;

  Query Match             22.0%;  Score 712;  DB 1;  Length 683;
  Best Local Similarity   30.2%;  Pred. No. 1.3e-41;
  Matches  188;  Conservative  121;  Mismatches  229;  Indels   84;  Gaps   20;

Qy         22 WEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVL----NLNAGEILQMFGKMFFVF 77
Db         73 WPNIHK-LKLDPQ-------TFKSCANYDYLADIQELLLKMDEASASEILVLLGEELITC 124

Qy         78 CQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDAEKGKGLILHYYS 137
Db        125 CCTGIIERAFRCLGTDLQEFLGSLDGVYDVLKLQEEDVTDTGFVCA----GEGELI-FTS 179

Qy        138 EREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC--------------DHTQFLIE 182
Db        180 ERPVIAWLLLGSLKALTRMLYKVDVNIKI--EPVERCPAVSLPLLAGQGQLPDHAD---G 234

Qy        183 EKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL-P 241
Db        235 SSTSVSKTIPETVQRSNSSNASDLQMNSSSFCKMFPWHFIMNEQLELVQLGRGFSKLYKP 294

Qy        242 QLQPGNCSLLSVFSLVRPH-IDISFHGILSHINTVFV--LRSKEGLLDVEKLECEDELTG 298
Db        295 YMADFGCQATTYFDFKRPKGLTMKFRDIVRRTYTPFLIGLNNPPGAVDFPAIG------- 347

Qy        299 TEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQ 358
Db        348 -----LEIKGQMVHCPESNSLLFIGSPFLDGLDGLTCNGLFISDIPLHDATREVILVGEQ 402

Qy        359 FREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYD 418
Db        403 ARAQDGLRRRMDKIKNSIEEANSAVTKERKKNVSLLHLIFPAEIAEKLWLGSSIDAKTYP 462

Qy        419 NVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPF--VYKVETVG 476
Db        463 DVTILFSDIVGFTSICSRAT----PFMVISMLEGLYKDFDEFCD-----FFDVYKVETIG 513

Qy        477 DKYMTVSGLPEPCIHHARSICHLALDMMEI----AGQVQVDGESVQITIGIHTGEVVTGV 532
Db        514 DAYCVASGLHRASIYDA----HRCLDGLKMIDACSKHITHDGEQIKMRIGLHTGTVLAGV 569

Qy        533 IGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSM 592
Db        570 VGRKMPRYCLFGHSVTIANKFESGSEALKINVSPTTKDWLTKHEGF--EFELQPRDP-SF 626

Qy        593 KGKKEPMQVWFLSRKNTGTEET 614
Db        627 LPKEFP---------NPGGTET 639
```

```
RESULT 12
ANPA_HUMAN
ID   ANPA_HUMAN     STANDARD;      PRT;  1061 AA.
AC   P16066;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)
DE   (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide
DE   A-type receptor).
GN   NPR1 OR ANPRA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=89356605; PubMed=2569967;
RA   Lowe D.G., Chang M.S., Hellmiss R., Chen E., Singh S., Garbers D.L.,
RA   Goeddel D.V.;
RT   "Human atrial natriuretic peptide receptor defines a new paradigm for
RT   second messenger signal transduction.";
RL   EMBO J. 8:1377-1384(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=98289600; PubMed=9618281;
RA   Takahashi Y., Nakayama T., Soma M., Izumi Y., Kanmatsuse K.;
RT   "Organization of the human natriuretic peptide receptor A gene.";
RL   Biochem. Biophys. Res. Commun. 246:736-739(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Maeda N., Knowles J.W.;
RT   "Identification of functional polymorphisms in noncoding regions of
RT   the human natriuretic peptide receptor A gene.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 634-1048 FROM N.A.
RC   TISSUE=Retina;
RX   MEDLINE=95042574; PubMed=7954658;
RA   Pardhasaradhi K., Kutty R.K., Gentleman S., Krishna G.;
RT   "Expression of mRNA for atrial natriuretic peptide receptor guanylate
RT   cyclase (ANPRA) in human retina.";
RL   Cell. Mol. Neurobiol. 14:1-7(1994).
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15357; CAA33417.1; -.
DR   EMBL; AB010491; BAA31199.1; -.
DR   EMBL; AB010470; BAA31199.1; JOINED.
DR   EMBL; AB010471; BAA31199.1; JOINED.
DR   EMBL; AB010472; BAA31199.1; JOINED.
DR   EMBL; AB010473; BAA31199.1; JOINED.
DR   EMBL; AB010474; BAA31199.1; JOINED.
DR   EMBL; AB010475; BAA31199.1; JOINED.
DR   EMBL; AB010476; BAA31199.1; JOINED.
DR   EMBL; AB010477; BAA31199.1; JOINED.
DR   EMBL; AB010478; BAA31199.1; JOINED.
DR   EMBL; AB010479; BAA31199.1; JOINED.
DR   EMBL; AB010480; BAA31199.1; JOINED.
DR   EMBL; AB010481; BAA31199.1; JOINED.
DR   EMBL; AB010482; BAA31199.1; JOINED.
DR   EMBL; AB010483; BAA31199.1; JOINED.
DR   EMBL; AB010484; BAA31199.1; JOINED.
DR   EMBL; AB010485; BAA31199.1; JOINED.
DR   EMBL; AB010486; BAA31199.1; JOINED.
DR   EMBL; AB010487; BAA31199.1; JOINED.
DR   EMBL; AB010488; BAA31199.1; JOINED.
DR   EMBL; AB010489; BAA31199.1; JOINED.
DR   EMBL; AB010490; BAA31199.1; JOINED.
DR   EMBL; AF190631; AAF01340.1; -.
DR   EMBL; S72628; AAD14112.1; -.
DR   PIR; S04459; OYHUAR.
DR   HSSP; Q02846; 1AWL.
DR   Genew; HGNC:7943; NPR1.
DR   MIM; 108960; -.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001170; Ntpep_receptorN.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL       1     32
FT   CHAIN       33   1061       ATRIAL NATRIURETIC PEPTIDE RECEPTOR A.
FT   DOMAIN      33    473       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   474    494       POTENTIAL.
FT   DOMAIN     495   1061       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     528    805       PROTEIN KINASE-LIKE.
FT   DOMAIN     876   1006       GUANYLATE CYCLASE.
FT   DISULFID    92    118       BY SIMILARITY.
FT   DISULFID   196    245       BY SIMILARITY.
FT   DISULFID   455    455       INTERCHAIN (PROBABLE).
FT   DISULFID   464    464       INTERCHAIN (PROBABLE).
FT   CARBOHYD    34     34       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    45     45       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   212    212       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   338    338       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   379    379       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   386    386       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   427    427       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1061 AA;  118918 MW;  E6B5BD0FCA32F70D CRC64;

  Query Match             14.9%;  Score 481.5;  DB 1;  Length 1061;
  Best Local Similarity   27.7%;  Pred. No. 1.8e-25;
  Matches  197;  Conservative   88;  Mismatches  227;  Indels  199;  Gaps   31;

Qy         33 EEGQFLVRIIYDDSKTYDLVAAASKVLN-------------------------LNAGEI 66
              | | | | : |: : : :|| : : ||                           |:  |:
Db        416 ENGAFRVVLNYNGT-SQELVAVSGRKLNWPLGYPPPDIPKCGFDNEDPACNQDHLSTLEV 474

Qy         67 LQMFGKM---------FFVF----CQESGYDTILRVLGSNVREFLQNLDALHDHLATIYP 113
              | : | :          ||::     ::     : ||   :|        :|  ||
Db        475 LALVGSLSLLGILIVSFFIYRKMQLEKELASELWRVRWEDVEP-----SSLERHL----- 524

Qy        114 GMRAPSFRCTDAEKGK--GLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRN 171
                |:   | | : :|   | :|      ||   :     ||   :  |  : :| : ::
Db        525 --RSAGSRLTLSGRGSNYGSLL----TTEGQFQV---FAKTAYYK--GNLVAVKRVNRKR 573

Qy        172 EECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVT 230
              |   |: :: | :   :   | | ||    |    |   | :|       | : : :
Db        574 IEL--TRKVLFELKHMRDVQNEHLTRFVGACTDPPNICILTEYCPRGSLQDILENESITL 631

Qy        231 ------QCGNAIYRVLPQLQPGN-CSLLSVFSLVRPHIDISFHGILSHINTV----FVLR 279
                   | | : :  |  |  ||                 || |   | |    |||:
Db        632 DWMFRYSLTNDIVKGMLFLHNGAICS---------------HGNLKSSNCVVDGRFVLK 675

Qy        280 SKE-GLLDVEKLECEDELTGTEISCLRLKGQMIY---LPEADSILFLCSPSV--MNLDDL 333
                : ||     |: |          :|  :|   |  |  :| : || |      |:
Db        676 ITDYGLESFRDLDPE------------QGHTVYAKKLWTAPELLRMASPPVRGSQAGDV 722

Qy        334 TRRGLYLSDIPL-----HDATRDL--------VLLGEQ--FREEYKLTQELEIL------ 372
                |: | :| |      |    ||        |  |||  ||    |   || |
```

```
Db        723 YSFGIILQEIALRSGVFHVEGLDLSPKEIIERVTRGEQPPFRPSLALQSHLEELGLLMQR 782

Qy        373 ----------TDRLQLTLR---------------------------------ALEDEK 387
                      :::||||                                    |  :||
Db        783 CWAEDPQERPPFQQIRLTLRKFNRENSSNILDNLLSRMEQYANNLEELVEERTQAYLEEK 842

Qy        388 KKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIV 447
              :| : ||| :|| ||| :|:    | |: :|:||| || |||| |  ::       |::|
Db        843 RKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAEST----PMQVV 898

Qy        448 NLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIH----HARSICHLALDM 503
               ||||||| || : |   |  ||||||:|| || |||||   :     ||  :  :|| :
Db        899 TLLNDLYTCFDAVID---NFDVYKVETIGDAYMVVSGLP---VRNGRLHACEVARMALAL 952

Qy        504 MEIAGQVQV---DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKG 560
              ::     ::     | ::: |||||| |  ||:| :||||||||:|||  || |: ||
Db        953 LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGEAL 1012

Qy        561 KINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSRKNTGT 611
              ||::|  |   |         | || || | |||| :   | |  : : |
Db       1013 KIHLSSETKAVL----EEFGGFELELRGDVEMKGKGKVRTYWLLGERGSST 1059
```

RESULT 13
ANPA_RAT
ID ANPA_RAT STANDARD; PRT; 1057 AA.
AC P18910;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)
DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide
DE A-type receptor).
GN NPR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=89143770; PubMed=2563900;
RA Chinkers M., Garbers D.L., Chang M.S., Lowe D.G., Chin H.,
RA Goeddel D.V., Schulz S.;
RT "A membrane form of guanylate cyclase is an atrial natriuretic
RT peptide receptor.";
RL Nature 338:78-83(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91056089; PubMed=1978722;
RA Yamaguchi M., Rutledge L.J., Garbers D.L.;
RT "The primary structure of the rat guanylyl cyclase A/atrial
RT natriuretic peptide receptor gene.";
RL J. Biol. Chem. 265:20414-20420(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91352095; PubMed=1679239;
RA Duda T., Goraczniak R.M., Sharma R.K.;
RT "Site-directed mutational analysis of a membrane guanylate cyclase
RT cDNA reveals the atrial natriuretic factor signaling site.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7882-7886(1991).
CC -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC CYCLASE ACTIVITY ON BINDING OF ANF.
CC -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC DOMAIN OF PROTEIN KINASES.
CC -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; X14773; CAA32881.1; -.
DR EMBL; J05677; AAA41200.1; -.
DR EMBL; M74535; AAA41202.1; -.
DR PIR; S03348; OYRTR.
DR HSSP; Q02846; 1AWL.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001054; G_cyclase.
DR InterPro; IPR001170; Ntpep_receptorN.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00211; guanylate_cyc; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00255; NATPEPTIDER.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00044; CYCc; 1.
DR PROSITE; PS00458; ANF_RECEPTORS; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW cGMP synthesis; Signal.

FT	SIGNAL	1	28
FT	CHAIN	29	1057 ATRIAL NATRIURETIC PEPTIDE RECEPTOR A.
FT	DOMAIN	29	469 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	470	490 POTENTIAL.
FT	DOMAIN	491	1057 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	524	801 PROTEIN KINASE-LIKE.
FT	DOMAIN	872	1002 GUANYLATE CYCLASE.
FT	DISULFID	88	114 BY SIMILARITY.
FT	DISULFID	192	241 BY SIMILARITY.
FT	DISULFID	451	451 INTERCHAIN (PROBABLE).
FT	DISULFID	460	460 INTERCHAIN (PROBABLE).
FT	CARBOHYD	41	41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	208	208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	334	334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	375	375 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	382	382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	423	423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	366	366 Q -> H (IN REF. 3).
FT	CONFLICT	392	392 L -> P (IN REF. 3).

SQ SEQUENCE 1057 AA; 118951 MW; 9EA9AE685AC05816 CRC64;

Query Match 14.5%; Score 468; DB 1; Length 1057;
Best Local Similarity 43.4%; Pred. No. 1.6e-24;
Matches 112; Conservative 38; Mismatches 82; Indels 26; Gaps 9;

```
Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
                |:|    | |:: :| | |   :||:| : ||| :|| ||| :|:.  | |: :|:|
Db        817 EQYANNLE-ELVEERTQAYL----EEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSV 871

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              || || |||| |  ::      |::| ||||||| || : |  | ||||||:|| ||
Db        872 TIYFSDIVGFTALSAEST----PMQVVTLLNDLYTCFDAVID---NFDVYKVETIGDAYM 924

Qy        481 TVSGLPEPCIH----HARSICHLALDMMEIAGQVQV---DGESVQITIGIHTGEVVTGVI 533
               |||||.  :     |||  :|| :::     ::     | ::: |||||| |  ||:
Db        925 VVSGLP---VRNGQLHAREVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVV 981

Qy        534 GQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMK 593
              | :|||||||:|||  || |: ||  ||::|  |   |  | | || || | ||
Db        982 GLKMPRYCLFGDTVNTASRMESNGEALKIHLSSETKAVL---EEFD-GFELELRGDVEMK 1037

Qy        594 GKKEPMQVWFLSRKNTGT 611
```

```
                 || :    | |  :   |
Db       1038 GKGKVRTYWLLGERGCST 1055


RESULT 14
ANPA_MOUSE
ID   ANPA_MOUSE     STANDARD;      PRT;  1057 AA.
AC   P18293;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)
DE   (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide
DE   A-type receptor).
GN   NPR1 OR NPRA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=90324219; PubMed=1973687;
RA   Pandey K.N., Singh S.;
RT   "Molecular cloning and expression of murine guanylate cyclase/atrial
RT   natriuretic factor receptor cDNA.";
RL   J. Biol. Chem. 265:12342-12348(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95139994; PubMed=7838126;
RA   Schoenfeld J.R., Sehl P., Quan C., Burnier J.P., Lowe D.G.;
RT   "Agonist selectivity for three species of natriuretic peptide
RT   receptor-A.";
RL   Mol. Pharmacol. 47:172-180(1995).
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05504; AAA37670.1; -.
DR   EMBL; L31932; AAA66945.1; -.
DR   PIR; A36568; OYMSAR.
DR   PIR; A33088; A33088.
DR   HSSP; Q02846; 1AWL.
DR   MGD; MGI:97371; Npr1.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001170; Ntpep_receptorN.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL       1     28
FT   CHAIN       29   1057       ATRIAL NATRIURETIC PEPTIDE RECEPTOR A.
FT   DOMAIN      29    469       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   470    490       POTENTIAL.
FT   DOMAIN     491   1057       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     524    801       PROTEIN KINASE-LIKE.
FT   DOMAIN     872   1002       GUANYLATE CYCLASE.
FT   DISULFID    88    114       BY SIMILARITY.
FT   DISULFID   192    241       BY SIMILARITY.
FT   DISULFID   451    451       INTERCHAIN (PROBABLE).
FT   DISULFID   460    460       INTERCHAIN (PROBABLE).
FT   CARBOHYD    41     41       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   334    334       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   375    375       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   382    382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     3      3       G -> R (IN REF. 1).
FT   CONFLICT    39     39       L -> V (IN REF. 1).
FT   CONFLICT   122    122       G -> D (IN REF. 1).
FT   CONFLICT   130    130       V -> L (IN REF. 1).
FT   CONFLICT   285    285       R -> E (IN REF. 1).
FT   CONFLICT   301    301       A -> R (IN REF. 1).
FT   CONFLICT   404    405       FS -> SP (IN REF. 1).
FT   CONFLICT   590    590       H -> Q (IN REF. 1).
FT   CONFLICT   652    652       G -> C (IN REF. 1).
FT   CONFLICT   833    833       A -> P (IN REF. 1).
FT   CONFLICT   958    958       R -> G (IN REF. 1).
FT   CONFLICT  1044   1044       T -> S (IN REF. 1).
FT   CONFLICT  1050   1050       E -> D (IN REF. 1).
FT   CONFLICT  1055   1057       TRG -> SRA (IN REF. 1).
SQ   SEQUENCE   1057 AA;  119109 MW;  53A544FB2C8EF253 CRC64;

  Query Match           14.4%;  Score 466;  DB 1;  Length 1057;
  Best Local Similarity  28.4%;  Pred. No. 2.1e-24;
  Matches  190;  Conservative   86;  Mismatches  210;  Indels  182;  Gaps   32;

Qy        25 IKKEAQLDEE-GQFLVRIIYDDSKTYDL---VAAASKVLNL-----NAGEILQMFGKMFF 75
             | :: ||::|    | |: ::| :   |   : :|   | |     | | :|   |: |
Db       489 IYRKMQLEKELVSELWRVRWEDLQPSSLERHLRSAGSRLTLSGRGSNYGSLLTTEGQ-FQ 547

Qy        76 VFCQESGYD----TILRVLGSNV---REFLQNL----DALHDHLATIYPGMRAPSFRCTD 124
             || : : |      : ||    :   |: |  |    |  ::|| | : |       |||
Db       548 VFAKTAYYKGNLVAVKRVNRKRIELTRKVLFELKHMRDVQNEHL-TRFVG------ACTD 600

Qy       125 AEKGKGLILHYYSEREGLQDIVIG---------------IIKTV------AQQIHG--- 159
                ||  |  |  ||||:                    |:| :      |  ||
Db       601 PP--NICILTEYCPRGSLQDILENESITLDWMFRYSLTNDIVKGMLFLHNGAIGSHGNLK 658

Qy       160 -------------TEIDMKVIQQRNEECDHTQFLIEEKESKEEDFYEDLDRFEE---NG 202
                         |:  ::  :    |  || |      :|:     :| |     . |
Db       659 SSNCVVDGRFVLKITDYGLESFRDPEPEQGHTLF------AKKLWTAPELLRMASPPARG 712

Qy       203 TQESRISPYTFCKAFPFHIIFDRDLVVTQCG-----------NAIYRVLPQLQPGNCSL 250
             :|   :      : | ||     :  : |              | ||    ||
Db       713 SQAGDV--------YSFGIILQE--IALRSGVFYVEGLDLSPKEIIERVTRGEQPP---- 758

Qy       251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310
                || :|:      ||:  :  |       :::   ||       :|| :
Db       759 ------FRPSMDLQ-----SHLEELGQL--------MQRCWAEDPQERPPFQQIRLALRK 799

Qy       311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370
               : :|          ||:|  |                        |:|    | |
Db       800 FNKENSSNI----------LDNLLSR----------------------MEQYANNLE-E 825

Qy       371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430
             :: :| |  |    :||:| : ||| :|| ||| :|:    | |: :|:||| || ||||
Db       826 LVEERTQAYL----EEKRKAEALLYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGF 881

Qy       431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490
```

```
Db        882 TALSAEST----PMQVVTLLNDLYTCFDAVID---NFDVYKVETIGDAYMVVSGLP---V 931

Qy        491 H----HARSICHLALDMMEIAGQVQV---DGESVQITIGIHTGEVVTGVIGQRMPRYCLF 543
Db        932 RNGQLHAREVARMALALLDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGVVGLKMPRYCLF 991

Qy        544 GNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWF 603
Db        992 GDTVNTASRMESNGEALRIHLSSETKAVL---EEFD-GFELELRGDVEMKGKGKVRTYWL 1047

Qy        604 LSRKNTGT 611
Db       1048 LGERGCST 1055


RESULT 15
ANPB_RAT
ID   ANPB_RAT       STANDARD;      PRT;  1047 AA.
AC   P16067;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B)
DE   (Guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide
DE   B-type receptor).
GN   NPR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89376566; PubMed=2570641;
RA   Schulz S., Singh S., Bellet R.A., Singh G., Tubb D.J., Chin H.,
RA   Garbers D.L.;
RT   "The primary structure of a plasma membrane guanylate cyclase
RT   demonstrates diversity within this new receptor family.";
RL   Cell 58:1155-1162(1989).
CC   -!- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC       CYCLASE ACTIVITY ON BINDING OF ANF. SEEMS TO BE STIMULATED MORE
CC       EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.
CC   -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC       WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC       WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC       CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC   -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC       DOMAIN OF PROTEIN KINASES.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M26896; AAA41205.1; -.
DR   PIR; A33300; OYRTBR.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001170; Ntpep_receptorN.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   PRINTS; PR00255; NATPEPTIDER.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00458; ANF_RECEPTORS; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW   cGMP synthesis; Signal.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23   1047       ATRIAL NATRIURETIC PEPTIDE RECEPTOR B.
FT   DOMAIN      23    458       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   459    478       POTENTIAL.
FT   DOMAIN     479   1047       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     513    786       PROTEIN KINASE-LIKE.
FT   DOMAIN     861    991       GUANYLATE CYCLASE.
FT   DISULFID    75    101       BY SIMILARITY.
FT   DISULFID   439    439       INTERCHAIN (PROBABLE).
FT   DISULFID   448    448       INTERCHAIN (PROBABLE).
FT   CARBOHYD    24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1047 AA;  117126 MW;  5062C49228CC14A3 CRC64;

  Query Match             14.1%;  Score 454.5;  DB 1;  Length 1047;
  Best Local Similarity   27.9%;  Pred. No. 1.3e-23;
  Matches  178;  Conservative   79;  Mismatches  238;  Indels  143;  Gaps   23;

Qy         58 VLNLNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVR-EFLQ--NLDALHDHLATIYPG 114
Db        458 IVALGTGITFIMFGVSSFLIFRKLMLEKELASMLWRIRWEELQFGNSDRYH--------- 508

Qy        115 MRAPSFRCTDAEKGKG--LILHYYSEREGLQDIVIGIIKTVAQQIH--GTEIDMKVIQQR 170
Db        509 -----------KGAGSRLTLSLRGSSYGSLMTAHGKYQIFANTGHFKGNVVAIKHVNKK 556

Qy        171 NEECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVT 230
Db        557 RIELTR-QVLFELKHMRDVQF-NHLTRF-----IGACIDPPNIC-----------IVTE 597

Qy        231 QCGNAIYRVLPQLQPGNCSLLSVFSLVRPHID---------ISFHGILSHINTV----FV 277
Db        598 YCPRGSLQDILENDSINLDWMFRYSLINDLVKGMAFLHNSIISSHGSLKSSNCVVDSRFV 657

Qy        278 LRSKE-GL--------------LDVEKLECEDEL--------TGTE----------ISCL 304
Db        658 LKITDYGLASFRSTAEPDDSHALYAKKLWTAPELLSGNPLPTTGMQKADVYSFAIILQEI 717

Qy        305 RLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGE------- 357
Db        718 ALRSGPFYLEGLD----LSPKEIVQKVRNGQRPYFRPSIDRTQLNEELVLLMERCWAQDP 773

Qy        358 --------------QFREE----------YKLTQELEILTDRLQLTLRALEDEKKKTDTL 393
Db        774 TERPDFGQIKGFIRRFNKEGGTSILDNLLLRMEQYANNLEKLVEERTQAYLEEKRKAEAL 833

Qy        394 LYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDL 453
Db        834 LYQILPHSVAEQLKRGETVQAEAFDSVTIYFSDIVGFTALSAEST----PMQVVTLLNDL 889

Qy        454 YTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLP-EPCIHHARSICHLALDMMEIAGQVQV 512
Db        890 YTCFDAIID---NFDVYKVETIGDAYMVVSGLPGRNGQRHAPEIARMALALLDAVSSFRI 946

Qy        513 ---DGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTY 569
Db        947 RHRPHDQLRLRIGVHTGPVCAGVVGLKMPRYCLFGDTVNTASRMESNGQALKIHVSSTTK 1006

Qy        570 RCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSRK 607
```

Db 1007 DAL----DELGCFQLELRGDVEMKGKGKMRTYWLLGER 1040

Search completed: June 27, 2003, 13:03:29
Job time : 12.9305 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:00:58 ; Search time 31.2101 Seconds
(without alignments)
4086.601 Million cell updates/sec

Title: US-09-762-767A-4
Perfect score: 3231
Sequence: 1 MYGFVNHALELLVIRNYGPE.....QVWFLSRKNTGTEETKQDDD 619

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3198	99.0	620	11	O54865	O54865 mus musculu
2	2887	89.4	614	13	P79998	P79998 oryzias lat
3	2859	88.5	617	13	Q90VY5	Q90vy5 fugu rubrip
4	1903.5	58.9	600	5	O77106	O77106 manduca sex
5	1883	58.3	758	5	Q24086	Q24086 drosophila
6	1883	58.3	787	5	Q9VA09	Q9va09 drosophila
7	1870.5	57.9	604	5	Q95NK5	Q95nk5 hemicentrot
8	1761	54.5	649	5	Q17010	Q17010 anopheles g
9	1105	34.2	742	11	Q91XJ7	Q91xj7 rattus norv
10	983	30.4	636	11	Q920Q1	Q920q1 rattus norv
11	960.5	29.7	685	5	Q9XTE0	Q9xte0 caenorhabdi
12	955	29.6	686	5	P92006	P92006 caenorhabdi
13	955	29.6	940	5	O76340	O76340 manduca sex
14	938	29.0	751	5	O02298	O02298 caenorhabdi
15	929.5	28.8	690	5	Q9VFC5	Q9vfc5 drosophila
16	914	28.3	699	5	Q17707	Q17707 caenorhabdi
17	912.5	28.2	684	5	Q9BI80	Q9bi80 caenorhabdi
18	788	24.4	564	4	Q9NNW8	Q9nnw8 homo sapien
19	788	24.4	675	13	Q90VV5	Q90vv5 fugu rubrip
20	786.5	24.3	690	4	Q8TAH3	Q8tah3 homo sapien
21	783	24.2	678	13	Q9PWI2	Q9pwi2 oryzias lat
22	778	24.1	678	13	P79997	P79997 oryzias lat
23	778	24.1	691	11	Q9DBQ3	Q9dbq3 mus musculu
24	778	24.1	691	11	Q9ERL9	Q9erl9 mus musculu
25	757	23.4	699	5	O77105	O77105 manduca sex
26	729.5	22.6	752	5	Q22301	Q22301 caenorhabdi
27	726	22.5	676	5	Q24085	Q24085 drosophila
28	720	22.3	676	5	Q95SQ4	Q95sq4 drosophila
29	704.5	21.8	669	5	Q9VEU5	Q9veu5 drosophila
30	701.5	21.7	583	5	O44468	O44468 caenorhabdi
31	685	21.2	649	5	Q9VEU6	Q9veu6 drosophila
32	646.5	20.0	947	5	P90895	P90895 caenorhabdi
33	565	17.5	108	6	Q8SPV3	Q8spv3 ovis aries
34	493	15.3	220	11	Q8R5L4	Q8r5l4 rattus norv
35	490.5	15.2	1070	13	O93490	O93490 anguilla ja
36	478.5	14.8	1126	5	Q9BPR0	Q9bpr0 bombyx mori
37	472	14.6	1055	13	Q98UI1	Q98ui1 oryzias lat
38	471	14.6	1055	13	Q9YGW3	Q9ygw3 oryzias lat
39	470	14.5	1057	5	Q18331	Q18331 caenorhabdi
40	469.5	14.5	1052	13	Q90YB7	Q90yb7 rana catesb
41	467.5	14.5	1276	5	Q8T8L6	Q8t8l6 caenorhabdi
42	466	14.4	1057	11	Q91X04	Q91x04 mus musculu
43	464	14.4	1047	13	Q9PWH0	Q9pwh0 xenopus lae
44	458	14.2	1076	5	Q9VF19	Q9vf19 drosophila
45	456.5	14.1	632	5	Q94429	Q94429 ciona intes

ALIGNMENTS

RESULT 1
O54865
ID O54865 PRELIMINARY; PRT; 620 AA.
AC O54865;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Soluble guanylate cyclase beta-1 subunit.
GN GUCY1B3 OR GC-S-BETA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Gansemans Y., Brouckaert P., Fiers W.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/10;
RX PubMed=10984516;
RA Sharina I.G., Krumenacker J.S., Martin E., Murad F.;
RT "Genomic organization of alpha 1 and beta 1 subunits of the mammalian
RT soluble guanylyl cyclase genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10878-10883(2000).
DR EMBL; AF020339; AAB94876.1; -.
DR EMBL; AF297083; AAG17447.1; -.
DR HSSP; P16068; 1AWN.
DR MGD; MGI:1860604; Gucy1b3.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYCc; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW Lyase.
SQ SEQUENCE 620 AA; 70597 MW; EED89ABF539F4EF9 CRC64;

Query Match 99.0%; Score 3198; DB 11; Length 620;

```
  Best Local Similarity   98.9%;  Pred. No. 3.1e-233;
  Matches  612;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRNLVVTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGAE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPLFHLEHRGPVSMKGKKEPMQ 600

Qy        601 VWFLSRKNTGTEETKQDDD 619
              ||||||||||||||  ::|:
Db        601 VWFLSRKNTGTEETNEEDE 619


RESULT 2
P79998
ID   P79998      PRELIMINARY;      PRT;   614 AA.
AC   P79998;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta subunit (EC 4.6.1.2).
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98237571; PubMed=9578459;
RA   Mikami T., Kusakabe T., Suzuki N.;
RT   "Molecular cloning of cDNAs and expression of mRNAs encoding alpha and
RT   beta subunits of soluble guanylyl cyclase from medaka fish Oryzias
RT   latipes.";
RL   Eur. J. Biochem. 253:42-48(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99303623; PubMed=10373466;
RA   Mikami T., Kusakabe T., Suzuki N.;
RT   "Tandem organization of medaka fish soluble guanylyl cyclase alpha1
RT   and beta1 subunit genes. Implications for coordinated transcription of
RT   two subunit genes.";
RL   J. Biol. Chem. 274:18567-18573(1999).
DR   EMBL; AB000850; BAA19199.1; -.
DR   EMBL; AB022281; BAA76691.1; -.
DR   HSSP; P16068; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   614 AA;  69917 MW;  64EDF9452D4714C1 CRC64;

  Query Match             89.4%;  Score 2887;  DB 13;  Length 614;
  Best Local Similarity   89.2%;  Pred. No. 9.8e-210;
  Matches  547;  Conservative   37;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              |||||||||||||:|||||||||||:|||| ||||||||||:|:|||||||||||||||
Db          1 MYGFVNHALELLVLRNYGPEVWEDIKREAQLDIEGQFLVRIIYEDAKTYDLVAAASKVLK 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              :|||:|||||||||| |||||||||||||||||||||||||||||| ||||||||||
Db         61 INAGDILQMFGKMFFEFCQESGYDTILRVLGSNVREFLQNLDALHDHLGTIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||||||||  ||||||||||||||||||||||||||||||:||||||::|||||  :||
Db        121 RCTDAEKGNNLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIEMKVIQQKSEECDHIKFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||:|:|| | |||| ||||||||:|||||||||||||::||:||::|||||||||||
Db        181 IEEKDSEEEAFNEDLDGFEENGTQETRISPYTFCKAFPFHLMFDKDLMLTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              ||||||:| | |||||||||| ||||||||||||||||||||:|| :| |||||| |
Db        241 PQLQPGSCILPSVFSLVRPHIDFSFHGILSHINTVFVLRSKEGLLNVETVENEDELTGVE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||::|||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEAENILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              |||||||||||||||| |||||||||||| |||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQHTLRALEDEKKKTDRLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              |||||||||| ||||||| |||:||||||||:||||| ||||||||:|||||||||||||
Db        421 TILFSGIVGFNTFCSKHASAEGAIKIVNLLNDVYTRFDILTDSRKNPYVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||||| |||:||||||||||||||||:|| ||||||||||||||||||||||||||
Db        481 TVSGLPEPCTHHAKSICHLALDMMEIAGQVKVDDESVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              |||||||||||||||||||||||||||||||||||    ||:|||| ||:|||::||||||:
Db        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLQCAENADPQFQLEYRGPITMKGKKEPMK 600

Qy        601 VWFLSRKNTGTEE 613
              |||||||::  ::
Db        601 VWFLSRKSSDADK 613


RESULT 3
Q90VY5
ID   Q90VY5      PRELIMINARY;      PRT;   617 AA.
AC   Q90VY5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta1 subunit.
```

```
GN   FRGCS-BETA1.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morinaga C., Yamamoto T., Moriya Y., Suzuki N.;
RT   "Identification of tandem organization of soluble guanylyl cyclase
RT   alpha1 and beta1 subunit genes in the Japanese pufferfish (Fugu
RT   rubripes) genome.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB062172; BAB60908.1; -.
DR   EMBL; AB062170; BAB60906.1; -.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR000834; Zn_carbOpept.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   PROSITE; PS00132; CARBOXYPEPT_ZN_1; UNKNOWN_1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   617 AA;  70236 MW;  7B531B5896A06191 CRC64;

  Query Match             88.5%;  Score 2859;  DB 13;  Length 617;
  Best Local Similarity   88.7%;  Pred. No. 1.3e-207;
  Matches  540;  Conservative   41;  Mismatches   28;  Indels    0;  Gaps    0;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||:|||||||||||||:|||||  ||||||||||:|:|||||||||||||
Db          1 MYGFVNHALELLVLRNYGPEVWEDIKREAQLDIEGQFLVRIIYEDAKTYDLVAAASKVLK 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ::||:|||:|||||| |||||||||||||||||||||||||||||||| ||||||||||
Db         61 IDAGDILQLFGKMFFEFCQESGYDTILRVLGSNVREFLQNLDALHDHLGTIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||||||||  |||||||||||||||||||||||||||||:||:|| :::|||| :||
Db        121 RCTDAEKGNSLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIEMKMIQPKSKECDHIKFL 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              ||||:|:|| |||||| ||||||||:||||||||||||::|||||::|||||||||||
Db        181 IEEKDSEEEAFYEDLDGFEENGTQETRISPYTFCKAFPFHLMFDRDLMLTQCGNAIYRVL 240

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              ||||||:| | |||||||||| ||||||||||||||||||||:|| :| |||||| |
Db        241 PQLQPGSCILPSVFSLVRPHIDFSFHGILSHINTVFVLRSKEGLLNVETVENEDELTGVE 300

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              ||||||||||||||::||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISCLRLKGQMIYLPEAENILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              |||||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||
Db        361 EEYKLTQELEILTDRLQHTLRALEDEKKKTDRLLYSVLPPSVANELRHKRPVPAKRYDNV 420

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              ||||||||||| ||||||| |||:|||||||||:||||| ||||| ||:||||||||||||
Db        421 TILFSGIVGFNTFCSKHASAEGAIKIVNLLNDVYTRFDILTDSRNNPYVYKVETVGDKYM 480

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              |||||||| |||:|||||||:|:||||||:|| : ||||||||||||||||||||||||
Db        481 TVSGLPEPCTHHAKSICHLALEMLEIAGQVKVDDKPVQITIGIHTGEVVTGVIGQRMPRY 540

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              ||||||||||||||||||:|||||:|:||| | ||:|||||||:||||:||||||||:
Db        541 CLFGNTVNLTSRTETTGEKGRINVSEFTFRCLQSAENADPQFHLEYRGPVTMKGKKEPMK 600

Qy        601 VWFLSRKNT 609
              ||||||| |
Db        601 VWFLSRKPT 609
```

```
RESULT 4
O77106
ID   O77106      PRELIMINARY;      PRT;   600 AA.
AC   O77106;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta-1 subunit.
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingiodea; Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98409773; PubMed=9736646;
RA   Nighorn A., Gibson N.J., Rivers D.M., Hildebrand J.G., Morton D.B.;
RT   "The nitric oxide-cGMP pathway may mediate communication between
RT   sensory afferents and projection neurons in the antennal lobe of
RT   manduca sexta.";
RL   J. Neurosci. 18:7244-7255(1998).
DR   EMBL; AF062751; AAC61264.1; -.
DR   HSSP; P16068; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   600 AA;  68099 MW;  7F7D0136E6ADACB9 CRC64;

  Query Match             58.9%;  Score 1903.5;  DB 5;  Length 600;
  Best Local Similarity   60.3%;  Pred. No. 2.1e-135;
  Matches  366;  Conservative   94;  Mismatches  138;  Indels    9;  Gaps    4;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||:||||||:: :  | || |||:| :  || |||| ||:|  ||:|: || :||
Db          1 MYGFVNYALELLVMKTFDEETWETIKKKADVAMEGSFLVRQIYEDEITYNLITAAVEVLQ 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              : |  ||::||| || |||:|||| ||:|||:  |:|||||| ||||| |:|||||:|||
Db         61 IPADAILELFGKTFFEFCQDSGYDKILQVLGATPRDFLQNLDGLHDHLGTLYPGMRSPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              |||:  :   |:|||||:| ||: |||||:|||| ::| ||: ::::: : ||||| |||
Db        121 RCTERPEDGALVLHYYSDRPGLEHIVIGIVKTVASKLHNTEVKVEILKTK-EECDHVQFL 179

Qy        181 IEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
              | | :       ::    |  : | ::|| |||: ||||::||||| : | |  : |:|
Db        180 ITETSTTGRVSAPEIAEI-ETLSLEPKVSPATFCRVFPFHLMFDRDLNIVQAGRTVSRLL 238

Qy        241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
              |::    | :  |   ||||::::|  :|:|||||:|||::|     |::   |     |
Db        239 PRVTRPGCKITDVLDTVRPHLEMTFANVLAHINTVYVLKTKP-----EEMSVTD--PHEE 291

Qy        301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
              |: |||||||:|:|| | ::| | ||| |||||||||| ::|||||||||||||: |||
Db        292 IASLRLKGQMLYIPETDVVVFQCYPSVTNLDDLTRRGLCIADIPLHDATRDLVLMSEQFE 351

Qy        361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
              :|||| ||:|||:|| | | || ||:||| |||||| ||| ||||:||||:||| |
Db        352 ADYKLTQNLEVLTDKLQQTFRELELEKQKTDRLLYSVLPISVATELRHRRPVPARRYDTV 411

Qy        421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
              |:|||||||  :|:::  :|||||| :|||||| || ||| ::|| ||||||||||||
Db        412 TLLFSGIVGFANYCARNSDHKGAMKIVRMLNDLYTAFDVLTDPKRNPNVYKVETVGDKYM 471

Qy        481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIGQRMPRY 540
              ||||||  : ||: |  |||||||::  | |||| | ||||||:||||||||| |||||
Db        472 AVSGLPEYEVAHAKHISLLALDMMDLSQTVTVDGEPVGITIGIHSGEVVTGVIGHRMPRY 531

Qy        541 CLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQ 600
              |||||||||||| ||||  | ||||| ||  ||  :| | || | :|| |:|||| ||||
```

```
Db        532 CLFGNTVNLTSRCETTGVPGTINVSEDTYNYLMREDNHDEQFELTYRGHVTMKGKAEPMQ 591

Qy        601 VWFLSRK 607
                  |||:||
Db        592 TWFLTRK 598


RESULT 5
Q24086
ID   Q24086      PRELIMINARY;      PRT;   758 AA.
AC   Q24086;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta subunit.
GN   GYC-BETA-100B OR DGCB1 OR CG1470.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=95318108; PubMed=7797526;
RA   Shah S., Hyde D.R.;
RT   "Two Drosophila genes that encode the alph and beta subunits of the
RT   brain soluble guanylyl cyclase.";
RL   J. Biol. Chem. 270:15368-15376(1995).
DR   EMBL; U27123; AAA87941.1; -.
DR   HSSP; P16068; 1AWN.
DR   FlyBase; FBgn0013973; Gyc-beta-100B.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   758 AA;  84444 MW;  FB1DC388142EEF09 CRC64;

  Query Match             58.3%;  Score 1883;  DB 5;  Length 758;
  Best Local Similarity   49.9%;  Pred. No. 1e-133;
  Matches  371;  Conservative  102;  Mismatches  125;  Indels  146;  Gaps    7;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||:|||||:::::| |:|| |||:| :  ||||||| ||||  ||:|: || ::||
Db          1 MYGFVNYALELLVLKHFGEEIWEKIKKKAMVSMEGQFLVRQIYDDEITYNLIGAAVEILN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              : | :||::||| || |||:|||| ||:|||:  |:||||||||||| |:|||||||
Db         61 IPADDILELFGKTFFEFCQDSGYDKILQVLGATPRDFLQNLDALHDHLGTLYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNE-------- 172
              |||  ||   |:|||||| ||: |||||:| || ::|| |::: ::::: |
Db        121 RCT--EKDGELLLHYYSERPGLEHIVIGIVKAVASKLHGVEVEIDIVKRKGEPIDEAEKE 178

Qy        173 ------------------------------------------------------------ 172

Db        179 RAIARENQQLLEDAVATTTTGSATVVLAPSSDAERNNNHNGSNGSNNNGMANNGNTVNVN 238

Qy        173 ----------------------------EC---------------------DHTQFLIE 182
                                          :|                     || ||||
Db        239 NNNDGQQIASETDPSIALSTCPIAQDSFDCDGDKEQKCLRLLKNKSDDIERYDHVQFLIR 298

Qy        183 E--------KESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGN 234
              |        ::|:::  :|::      :   ||| |||| ||||::||| : : | |
Db        299 EINVAAKSQVDAKKDEVPDDMEFL----CEAPLISPATFCKVFPFHLMFDRQMKIVQAGK 354

Qy        235 AIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECED 294
              |: ||:|::   ||||: |   :|||: ::|  |||||||::||::::| :
Db        355 AVSRVIPRVAEENCSLIEVVEAIRPHLQLNFENILSHINTIYVLQTRQGAMS-------- 406

Qy        295 ELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVL 354
                                                                  :  |   |||||||:|:|| | ||| | ||||||||||::|||:|:|||||| |||||
Db        407 --SRHEQRFLRLKGQMMYIPETDRILFQCYPSVMNLDDLTKKGLYISDVPLHDAARDLVL 464

Qy        355 LGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPA 414
              | |:|  |||||: ||:|||:|| | | || ||:||| ||||||| ||||||||:||||
Db        465 LSEKFEAEYKLTKNLEMLTDKLQQTFRDLESEKQKTDRLLYSVLPKSVANELRHQRPVPP 524

Qy        415 KRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVET 474
              ||||:||::|||||||  :|: :   :|||||| :|||:||| || ||||::|  ||||||
Db        525 KRYDSVTLMFSGIVGFGQYCAANTDPDGAMKIVKMLNELYTVFDALTDSKRNLNVYKVET 584

Qy        475 VGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIG 534
              |||||| |||||: |  ||: :  :|||||::|  |::    ||||||||:|||||||||
Db        585 VGDKYMAVSGLPDHCEDHAKCMARVALDMMDMAKNVKMGSNPVQITIGIHSGEVVTGVIG 644

Qy        535 QRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKG 594
              |:|||||||||||||||||||  |:|||| ||| |    | |  ||||:|||| |||
Db        645 NRVPRYCLFGNTVNLTSRTETTGVPGRINVSEETYRLLCMAINQDDSFHLEYRGPVIMKG 704

Qy        595 KKEPMQVWFLSRKNT----GTEET 614
              |  ||  |||:|  :    ||  |
Db        705 KPTPMDCWFLTRATSSILGGTSST 728


RESULT 6
Q9VA09
ID   Q9VA09      PRELIMINARY;      PRT;   787 AA.
AC   Q9VA09;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   GYCBETA100B protein.
GN   GYC-BETA-100B OR GYCBETA100B OR CG1470.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
```

```
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003777; AAF57119.1; -.
DR   HSSP; P16068; 1AWN.
DR   FlyBase; FBgn0013973; Gyc-beta-100B.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase.
SQ   SEQUENCE   787 AA;  86849 MW;  ABD2F4E6D09C72B8 CRC64;

  Query Match             58.3%;  Score 1883;  DB 5;  Length 787;
  Best Local Similarity   49.9%;  Pred. No. 1.1e-133;
  Matches  371;  Conservative  102;  Mismatches  125;  Indels  146;  Gaps    7;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||:||||||:::::| |:|| |||:| :  ||||||| ||||  ||:|: || ::||
Db          1 MYGFVNYALELLVLKHFGEEIWEKIKKKAMVSMEGQFLVRQIYDDEITYNLIGAAVEILN 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              : | :|||::||| || |||:|||| ||:|||:  |:||||||||||| |:|||||||||
Db         61 IPADDILELFGKTFFEFCQDSGYDKILQVLGATPRDFLQNLDALHDHLGTLYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNE-------- 172
              |||  ||   |:|||||| ||: |||||:| || ::|| |::: ::::: |
Db        121 RCT--EKDGELLLHYYSERPGLEHIVIGIVKAVASKLHGVEVEIDIVKRKGEPIDEAEKE 178

Qy        173 ------------------------------------------------------------ 172

Db        179 RAIARENQQLLEDAVATTTTGSATVVLAPSSDAERNNNHNGSNGSNNNGMANNGNTVNVN 238

Qy        173 ----------------------------EC--------------------DHTQFLIE 182
                                           :|                    || ||||
Db        239 NNNDGQQIASETDPSIALSTCPIAQDSFDCDGDKEQKCLRLLKNKSDDIERYDHVQFLIR 298

Qy        183 E--------KESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGN 234
              |         ::|:::  :|::       :   ||| |||| ||||::||| : : | |
Db        299 EINVAAKSQVDAKKDEVPDDMEFL----CEAPLISPATFCKVFPFHLMFDRQMKIVQAGK 354

Qy        235 AIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECED 294
              |: ||:|::   ||||: |   :|||: ::|  |||||||::||::::| :
Db        355 AVSRVIPRVAEENCSLIEVVEAIRPHLQLNFENILSHINTIYVLQTRQGAMS-------- 406

Qy        295 ELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVL 354
                :  |   |||||||:|:|| | ||| | |||||||||::|||:||:|||| |||||
Db        407 --SRHEQRFLRLKGQMMYIPETDRILFQCYPSVMNLDDLTKKGLYISDVPLHDAARDLVL 464

Qy        355 LGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPA 414
              | |:|  |||||: ||:|||:|| | | || ||:||| ||||||| ||||||||:||||
Db        465 LSEKFEAEYKLTKNLEMLTDKLQQTFRDLESEKQKTDRLLYSVLPKSVANELRHQRPVPP 524

Qy        415 KRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVET 474
              ||||:||::|||||||  :|: :   :|||||| :||:||| || ||||::|  ||||||
Db        525 KRYDSVTLMFSGIVGFGQYCAANTDPDGAMKIVKMLNELYTVFDALTDSKRNLNVYKVET 584

Qy        475 VGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVIG 534
              |||||| |||||: |  ||: :  :|||||::|  |::    |||||||:|||||||||
Db        585 VGDKYMAVSGLPDHCEDHAKCMARVALDMMDMAKNVKMGSNPVQITIGIHSGEVVTGVIG 644

Qy        535 QRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKG 594
              |:|||||||||||||||||||  |:|||||  ||| |    | |  ||||:|||| |||
Db        645 NRVPRYCLFGNTVNLTSRTETTGVPGRINVSEETYRLLCMAINQDDSFHLEYRGPVIMKG 704

Qy        595 KKEPMQVWFLSRKNT----GTEET 614
              |  ||  |||:|  :    || |
Db        705 KPTPMDCWFLTRATSSILGGTSST 728
```

```
RESULT 7
Q95NK5
ID   Q95NK5      PRELIMINARY;      PRT;   604 AA.
AC   Q95NK5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta1.
GN   HPGCS-BETA1.
OS   Hemicentrotus pulcherrimus (Sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Hemicentrotus.
OX   NCBI_TaxID=7650;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Suzuki N.;
RT   "Genomic Structure and Expression of the Sea Urchin Soluble Guanylyl
RT   Cyclase beta Subunit Genes.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB062390; BAB58877.1; -.
DR   EMBL; AB062387; BAB58877.1; JOINED.
DR   EMBL; AB062388; BAB58877.1; JOINED.
DR   EMBL; AB062389; BAB58877.1; JOINED.
DR   EMBL; AB062386; BAB56135.1; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   604 AA;  68198 MW;  B1C79618705934D2 CRC64;

  Query Match             57.9%;  Score 1870.5;  DB 5;  Length 604;
  Best Local Similarity   57.2%;  Pred. No. 6.7e-133;
  Matches  351;  Conservative  112;  Mismatches  126;  Indels   25;  Gaps    5;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||||||||||||:| :| : ||:||:|| :: || |||||:|||  :|||| || |||
Db          1 MYGFVNHALELLVLREHGKDKWEEIKREAAVEIEGSFLVRIVYDDVLSYDLVGAAVKVLE 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ::| ::|: ||:||| || ||||| || ||||  | ||||||||||||:|||||||||||
Db         61 ISANDLLEAFGRMFFEFCVESGYDNILNVLGSTTRHFLQNLDALHDHLASIYPGMRAPSF 120

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQFL 180
              ||:  |    |:|||||||| ||: ||||::::||: :||:|: :::|: : |:||| ||
Db        121 RCSTRESDGALVLHYYSERPGLEHIVIGLVRSVAKTLHGSEVHVEIIKNKGEDCDHVQFA 180

Qy        181 IEEKESKEEDFYEDLDRFEENGTQ-------ESRISPYTFCKAFPFHIIFDRDLVVTQCG 233
              | ||         :  : |:   |       | :||| | |:  ||||:|: :| | | |
Db        181 IIEK--------VETAKIEKQARQNLLALSKEPKISPSTLCRILPFHIMFNAELNVEQAG 232

Qy        234 NAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECE 293
              |:| |::| :   || :  :| :||||:: :|  |||| ||::||::  |:::
Db        233 NSIQRIVPNIINPNCRMTDLFHIVRPHMEFTFKSILSHANTIYVLKTNSGVVN------P 286

Qy        294 DELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLV 353
              :    |  |:|||||:::||:: :|:|||| |:|||:| :| ||||||||||||||||
Db        287 NNPRNGSIPALKLKGQMLHVPESNVLLYLCSPHVINLDELRQRELYLSDIPLHDATRDLV 346

Qy        354 LLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVP 413
              |: |:| |||||||:|||||:|| | | :|:|||||| ||||:|||||||||| ||||
Db        347 LISERFDEEYKLTQKLEILTDKLQQTYREIENEKKKTDRLLYSILPPSVANELRHHRPVP 406

Qy        414 AKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVE 473
              ||::: ||::|||| ||  || :::    |||||:||| :||:|| | ::  || |||||
Db        407 AKKFECVTLMFSGIFGFGDFCRRYS--HDAMKIVSLLNSVYTKFDVLMEN--NPDVYKVE 462

Qy        474 TVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQVDGESVQITIGIHTGEVVTGVI 533
              ||||||| ||||| ||  ||: |  :||:| |::  | ::|: : ||||:::|||||||:
Db        463 TVGDKYMAVSGLPVPCADHAKCIAKMALEMKELSADVIMEGDPIVITIGVYSGEVVTGVV 522
```

```
Qy        534 GQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMK 593
              ||||||||||||||||||||||||  ||||:::  | ||| |:|:|| |  : || |:||
Db        523 GQRMPRYCLFGNTVNLTSRTETTGVTGKINIADTAYDCLMEPQNADPTFQFDFRGLVNMK 582

Qy        594 GKKEPMQVWFLSRK 607
              || :|   : ||||.
Db        583 GKPKPCPCYLLSRK 596


RESULT 8
Q17010
ID   Q17010      PRELIMINARY;      PRT;   649 AA.
AC   Q17010; Q17007; Q17008; Q17009;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Soluble guanylyl cyclase beta subunit (Fragment).
GN   GCSBETA.
OS   Anopheles gambiae (African malaria mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC   Anopheles.
OX   NCBI_TaxID=7165;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SUA;
RX   MEDLINE=99124361; PubMed=9927171;
RA   Caccone A., Garcia B.A., Mathiopoulos K.D., Min G.S., Moriyama E.N.,
RA   Powell J.R.;
RT   "Characterization of the soluble guanylyl cyclase beta-subunit gene in
RT   the mosquito Anopheles gambiae.";
RL   Insect Mol. Biol. 8:23-30(1999).
RN   [2]
RP   SEQUENCE OF 539-649 FROM N.A.
RC   STRAIN=SUA, G3, GMMKG, AND MU;
RX   MEDLINE=96400927; PubMed=8807303;
RA   Garcia B.A., Caccone A., Mathiopoulos K.D., Powell J.R.;
RT   "Inversion monophyly in African anopheline malaria vectors.";
RL   Genetics 143:1313-1320(1996).
DR   EMBL; U42613; AAC47142.1; -.
DR   EMBL; AF017062; AAC47144.2; -.
DR   EMBL; U42612; AAC47141.1; -.
DR   EMBL; U42614; AAC47143.1; -.
DR   HSSP; P16068; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 2.
KW   Lyase.
FT   NON_TER       1      1
SQ   SEQUENCE   649 AA;  72514 MW;  376BAC40422C59DB CRC64;

  Query Match           54.5%;  Score 1761;  DB 5;  Length 649;
  Best Local Similarity   52.0%;  Pred. No. 1.4e-124;
  Matches  345;  Conservative   94;  Mismatches  116;  Indels  108;  Gaps    4;

Qy          2 YGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNL 61
              |||||:||||||::|:|  :|| |||:||:: ||||||| ||:|  ||:|: ||  :||:
Db          1 YGFVNYALELLVLKNFGLNIWEQIKKKAQVNMEGQFLVRQIYEDDITYNLIEAAVDILNI 60

Qy         62 NAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFR 121
               ||:||::||| || |||:|||| ||:|||:  |:|||||||||||| |:||||||||||
Db         61 PAGDILELFGKTFFEFCQDSGYDKILQVLGATPRDFLQNLDALHDHLGTLYPGMRAPSFR 120

Qy        122 CTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRN---------- 171
              ||  |    |:|||||| ||: |||||:| || ::|| ::::|:|:::
Db        121 CT--ETNGQLVLHYYSERPGLEHIVIGIVKAVASKLHGVDVEIKIIRRKGDPVEPEAKKA 178

Qy        172 ------------------------------------------------------------ 171

Db        179 TAEPPVANTAAAATPKARHSIPEVVKSVPITSLDPAVPELANLGLCKRILASKTSSSGGP 238

Qy        172 --------------------EECDHTQFLIEE----KESKEEDFYEDLDRFEENGTQESR 207
                                  |  || |||| |    |        :|     :   :|
Db        239 ARNGPEWQQKGPKGPGTTAVERSDHFQFLITEISGPKTPTRRSDEKDQQAVCQLVAKEPM 298

Qy        208 ISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHG 267
              ||| |||| ||||::|||:: : | | :: ||:|::   || ||::|  ||||: :||
Db        299 ISPMTFCKIFPFHLMFDRNMHIVQAGRSVSRVIPRIYEKNCPLLALFEAVRPHLQLSFEN 358

Qy        268 ILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSV 327
              ||:||||::||::| |::   :              |||||||:|:| :| ||| | |||
Db        359 ILAHINTIYVLKTKAGVMSKSE------------RYLRLKGQMMYIPGSDLILFQCYPSV 406

Qy        328 MNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEK 387
              |||||||::||::|||||||:|||||| |:|  |||||  |||||||| | | || ||
Db        407 MNLDDLTKKGLHISDIPLHDASRDLVLLSEKFEAEYKLTTNLEILTDRLQQTYRDLESEK 466

Qy        388 KKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIV 447
              :||| ||||||| :|||||||:|||  ||||:||::|||||||  :|: :   |||||||
Db        467 QKTDRLLYSVLPKTVANELRHQRPVAPKRYDSVTLMFSGIVGFGQYCAANTDPEGAMKIV 526

Qy        448 NLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIA 507
              :||:||| || |||||: |  :|||||||||| |||||:| :||: |  ||||:::|
Db        527 KMLNELYTIFDELTDSKSNSNIYKVETVGDKYMAVSGLPDECENHAKCIARLALDMLDMA 586

Qy        508 GQVQVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEY 567
              | :  |:::||||||:|||||||| ||||||||||||||||||||  | ||:||
Db        587 KNVMMGTEAMKITIGIHSGEVVTGVIGNRMPRYCLFGNTVNLTSRTETTGVPGHINISET 646

Qy        568 TYR 570
              |||
Db        647 TYK 649


RESULT 9
Q91XJ7
ID   Q91XJ7      PRELIMINARY;      PRT;   742 AA.
AC   Q91XJ7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylyl cyclase beta 2 subunit.
GN   GUCY1B2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX   MEDLINE=21391941; PubMed=11406623;
RA   Koglin M., Vehse K., Budaeus L., Scholz H., Behrends S.;
RT   "Nitric Oxide Activates the beta 2 Subunit of Soluble Guanylyl Cyclase
RT   in the Absence of a Second Subunit.";
RL   J. Biol. Chem. 276:30737-30743(2001).
DR   EMBL; AY004153; AAF86581.1; -.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ   SEQUENCE   742 AA;  83050 MW;  41CE148B15E244E4 CRC64;

  Query Match           34.2%;  Score 1105;  DB 11;  Length 742;
  Best Local Similarity   36.9%;  Pred. No. 6.6e-75;
  Matches  249;  Conservative  117;  Mismatches  221;  Indels   88;  Gaps   13;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||||:|  |: ||   :| | || :|  |:: :   |:   :|||  |  |:  | |||:
Db          1 MYGFINTCLQSLVTEKFGEETWEKLKASAEVQD--AFMTYTVYDDIITIKLIQEACKVLD 58

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
               ::   ||::||: || ||: |||| :|| || |: ||::|||||| :||  |  | ||||
```

```
Db        59 VSMEAILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDALHSYLALSYQEMNAPSF 118

Qy       121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC------ 174
             |  :   |  ::|||||:| ||  || |||: ||:    |:: | ::   |||
Db       119 RVEEGADG-AMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMSILDM-NEEVERTGKK 176

Qy       175 DHTQFLIEEKESKE-----------EDFYED---------------------------- 194
             :|  ||: :|  ::           ||   |
Db       177 EHVVFLVVQKAHRQIRGAKASRPQGSEDSQADQEALQGTLLRMKERYLNIPVCPGEKSHS 236

Qy       195 -------------LDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVL 240
                         | |:    :   :    || ||||||:||  | | | |  | : :
Db       237 TAVRASVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFHIVFDEALRVKQAGVNIQKYV 296

Qy       241 PQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTE 300
             | :    :|   ||:: | :  :   |   ||: |||:::: ::   :
Db       297 PGILTQKFALDEYFSIIHPQVTFNISSICKFINSQFVLKTRKEMMPKAR---------KS 347

Qy       301 ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFR 360
                |:|:||||::     ::|:|||:| :| :|    ::|||  || ||||:|| :|
Db       348 QPMLKLRGQMIWMESLRCMIFMCSPNVRSLQELEESKMHLSDIAPHDTTRDLILLNQQRL 407

Qy       361 EEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNV 420
             | :|: :||   : |::    |  |||||:||||::||  |||:|:  | | |  ::
Db       408 AEMELSCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVANQLKEGRKVAAGEFETC 467

Qy       421 TILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYM 480
             ||||| :| |   |   |: |  ::|||:|| :|::|| ||       ||||||:|| ||
Db       468 TILFSDVVTFTNIC---AACE-PIQIVNMLNSMYSKFDRLTSVHD---VYKVETIGDAYM 520

Qy       481 TVSGLPEPCIHHARSICHLALDMMEIAGQVQ--VDGESVQITIGIHTGEVVTGVIGQRMP 538
             | |:| |   ||: : : || |   | :|   | || :|| :||||| |: ||:| :||
Db       521 VVGGVPVPVESHAQRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVVGDKMP 580

Qy       539 RYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEP 598
             ||||||:|||  || |: |   |:::|   :| |      :  | :  || : :||| :
Db       581 RYCLFGDTVNTASRMESHGLPSKVHLSPTAHRAL-----KNKGFEIVRRGEIEVKGKGK- 634

Qy       599 MQVWFLSRKNTGTEE 613
             |  :|| :    ||:
Db       635 MTTYFLIQNLNATED 649


RESULT 10
Q920Q1
ID   Q920Q1      PRELIMINARY;      PRT;   636 AA.
AC   Q920Q1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Soluble guanylate cyclase beta 2b.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA   Okamoto H., Asakawa T.;
RT   "cDNA cloning and expression analysis of novel variant of rat sGC
RT   beta2 subunit.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB058888; BAB68564.1; -.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   636 AA;  71642 MW;  951DBDE8970E0890 CRC64;

  Query Match             30.4%;  Score 983;  DB 11;  Length 636;
  Best Local Similarity   36.7%;  Pred. No. 8.9e-66;
  Matches  224;  Conservative  104;  Mismatches  196;  Indels   86;  Gaps   12;

Qy        66 ILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDA 125
             ||:::|: || ||: |||| :|| || |: ||::|||||| :|| _|  | |||||  :
Db         4 ILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDALHSYLALSYQEMNAPSFRVEEG 63

Qy       126 EKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC------DHTQF 179
               |  ::|||||:| ||  || |||: ||:    |:: | ::   |||       :|  |
Db        64 ADG-AMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMSILDM-NEEVERTGKKEHVVF 121

Qy       180 LIEEKESKE-----------EDFYED------------------------------ 194
             |: :|  ::            ||  |
Db       122 LVVQKAHRQIRGAKASRPQGSEDSQADQEALQGTLLRMKERYLNIPVCPGEKSHSTAVRA 181

Qy       195 ---------LDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQP 245
                       | |:    :   :    || ||||||:||  | | | |  | : :| :
Db       182 SVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFHIVFDEALRVKQAGVNIQKYVPGILT 241

Qy       246 GNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLR 305
               :|   ||:: | :  :   |   ||: |||:::: ::   :              |:
Db       242 QKFALDEYFSIIHPQVTFNISSICKFINSQFVLKTRKEMMPKAR---------KSQPMLK 292

Qy       306 LKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKL 365
             |:||||::     ::|:||| : :| :|    ::||||  || ||||:|| :|   | :|
Db       293 LRGQMIWMESLRCMIFMCSPKLRSLQELEESKMHLSDIAPHDTTRDLILLNQQRLAEMEL 352

Qy       366 TQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFS 425
             : :||   : |::    |  |||||:||||::||  |||:|:  | | |  ::  |||||
Db       353 SCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVANQLKEGRKVAAGEFETCTILFS 412

Qy       426 GIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGL 485
               :| |   |   |: |  ::|||:|| :|::|| ||       ||||||:|| || | |:
Db       413 DVVTFTNIC---AACE-PIQIVNMLNSMYSKFDRLTSVHD---VYKVETIGDAYMVVGGV 465

Qy       486 PEPCIHHARSICHLALDMMEIAGQVQ--VDGESVQITIGIHTGEVVTGVIGQRMPRYCLF 543
             | |   ||: : : || |   | :|   | || :|| :||||| |: ||:| :|||||||
Db       466 PVPVESHAQRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVVGDKMPRYCLF 525

Qy       544 GNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWF 603
             |:|||  || |: |   |:::|   :| |      :  | :  || : :||| : |  :|
Db       526 GDTVNTASRMESHGLPSKVHLSPTAHRAL-----KNKGFEIVRRGEIEVKGKGK-MTTYF 579

Qy       604 LSRKNTGTEE 613
             | :    ||:
Db       580 LIQNLNATED 589


RESULT 11
Q9XTE0
ID   Q9XTE0      PRELIMINARY;      PRT;   685 AA.
AC   Q9XTE0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   C46E1.2 protein.
GN   C46E1.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lennard N.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
```

```
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lennard N.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z99942; CAB17073.1; -.
DR   EMBL; AL008867; CAB17073.1; JOINED.
DR   EMBL; AL008867; CAA15512.1; -.
DR   EMBL; Z99942; CAA15512.1; JOINED.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   Lyase.
SQ   SEQUENCE   685 AA;  77871 MW;  68064419692A7BBE CRC64;

  Query Match             29.7%;  Score 960.5;  DB 5;  Length 685;
  Best Local Similarity   34.2%;  Pred. No. 4.9e-64;
  Matches  220;  Conservative  131;  Mismatches  234;  Indels   59;  Gaps   17;

Qy          2 YGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNL 61
              :||:: ::  |:|| ||   |  : : |  :   : ::   | |: |: || | | :| :
Db         12 FGFIHESIRQLMIRTYGEAFWSKVLERAGFEAGKENIINHYYSDADTFSLVDAVSVILKV 71

Qy         62 NAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALH---DHLATIY-PGMRA 117
                :: :|:|     :  |:|:| ::| :  |:: || |||:||   ||:  :|   :|
Db         72 TREQVWEMYGCFLIQYTMETGWDDLIRSMSPNLKGFLDNLDSLHYFIDHV--VYKANLRG 129

Qy        118 PSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVI--QQRNEEC- 174
              ||||| |   |  : ||||: | ||  || |:::  |:::   :: | :    ||: :
Db        130 PSFRCEDNPDGT-ITLHYYTGRPGLYPIVKGVLREAAKRVFKLDVSMTITGRTQRSVQMA 188

Qy        175 ------DHTQFLIE---EKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDR 225
                   :|  ||::     :| ||     : :   |   : |::   |   ||:|:: |:
Db        189 TGERIEEHVIFLVKTLNTDQSNEEALGTAVVQHSNN--YKIRLTHMDFISTFPYHMVVDQ 246

Qy        226 DLVVTQCGNAIYRVLPQ-LQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGL 284
              |  : | |  :|  :|: |       |: :| : || | : |  | : || ||||: |
Db        247 DCKIVQVGRELYNHIPKDLLSVGTPLMRIFEVTRPQIPLDFDSICNFINAVFVLQVKTTP 306

Qy        285 LDVEK---------LECEDE--------LTGTEISCLRLKGQMIYLPEADSILFLCSPSV 327
              :: ::          :|  :          |  ::   |:|||||: :      |::|||| |
Db        307 MEFQRNANKRAAQAIEASENLYEDNNGALALSQSQHLKLKGQMMLMSSGGHIMYLCSPYV 366

Qy        328 MNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEK 387
              :: :| : || |: :|:|| ||||:|| :|   : ::  :||   ::|:   : || ||
Db        367 TSIPELLQYGLRLTAMPIHDPTRDLILLNQQRLSDVEMNLQLEANNEQLENMAKDLEVEK 426

Qy        388 KKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFN---AFCSKHASGEGAM 444
              |||  ||  :|||||| :|:    | |: |:  |::|: :  |      |:
Db        427 GKTDALLREMLPPSVAQQLKQGLSVEAREYEEATVMFTDVPTFQQIVPLCT-------PK 479

Qy        445 KIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMM 504
              ||:|||:|:|:|| |   :|    |||||||| ||:| |:|:    |  ||||||  |:
Db        480 DIVHLLNELFTKFDRLIGIQK---AYKVETVGDSYMSVGGIPDLVDDHCEVICHLALGMV 536

Qy        505 EIAGQV--QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKI 562
                |  |   :   : |  |||:| || ||:| :|||||||:|||  :|| |:    |:|
Db        537 MEARTVCDPITNTPLHIRAGIHSGPVVAGVVGAKMPRYCLFGDTVNTSSRMESHSPIGRI 596

Qy        563 NVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSR 606
              : ||   :|      |  :|  | || | :||| | |  :|| |
Db        597 HCSENAKKC----AESTGRFEFEPRGRVQIKGKGE-MNTYFLLR 635
```

```
RESULT 12
P92006
ID   P92006      PRELIMINARY;      PRT;   686 AA.
AC   P92006;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   M04G12.3 protein.
GN   M04G12.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sims M.A.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81103; CAB03210.1; -.
DR   HSSP; P16068; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ   SEQUENCE   686 AA;  78588 MW;  65AE352DB53CBA57 CRC64;

  Query Match             29.6%;  Score 955;  DB 5;  Length 686;
  Best Local Similarity   34.2%;  Pred. No. 1.3e-63;
  Matches  230;  Conservative  126;  Mismatches  235;  Indels   82;  Gaps   20;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              |:||:: ::  |||| || :||  : : :  :   : :|   | |: || || : | ||
Db          1 MFGFIHESIRQLVIRKYGEDVWLQVLERSGFENGKENIVNHYYSDTDTYVLVDSVSIVLK 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALH---DHLATIY-PGMR 116
                :  :| :|:|     :  | |:| ::| :  |:: || |||:||   ||:  :|   :|
Db         61 VTKDQIWEMYGGFLITYSMEIGWDELVRSMSPNLKGFLDNLDSLHYFIDHV--VYKANLR 118

Qy        117 APSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC-- 174
              |||||  :   |  |:|||:: | ||  || |::| ||: :   :| : |:| | :
Db        119 GPSFRCEENPDGT-LMLHYFTGRPGLYHIVKGVVKEVAKLVFNLDISL-VVQGRTQRSVH 176

Qy        175 --------DHTQFLIEEKE--------------SKEEDFYEDLDRFEENGTQESRISPY 211
                     :|  |||:  |                | | || | :|          ::|
Db        177 MNNGERVEEHVIFLIKNVEEPRRDSDTSTTSALTSVEPDFGEIID-------DNLKVSLQ 229

Qy        212 TFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQ--LQPGNCSLLSVFSLVRPHIDISFHGIL 269
              | :| |:| : |    : |||: :|  :|   ||||   :| :| : || | :  |  |
Db        230 DFSRALPYHFVLDESCRLVQCGDELYNHIPNELLQPGT-PILRIFEINRPQIPLDFENIC 288

Qy        270 SHINTVFVLRSKEGLL--------------DVEKLECEDELTGTEISC-LRLKGQMIYL 313
              : || ||||: |   |               :|| :| | |        | |:|||||: |
Db        289 NFINAVFVLQVKTSPLRKKHMNAMTKEEREQEVEAMEEEVESNELTQGCHLKLKGQMMML 348

Qy        314 PEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELEILT 373
                    |::|||| | ::::| : |: |: :|||||||||:|| :|   : ::  :||
Db        349 STKKHIIYLCSPYVTSINELMQFGMRLTAMPLHDATRDLILLNQQRLSDVEVNLQLEANN 408

Qy        374 DRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGFNAF 433
              ::|:     || |::|||::|  :||  :| :|     :    |:  |::|  :  |
Db        409 EQLETMTHELEVERQKTDSILKDMLPRKIAKQLLSGEHLEPCEYE-ATVMFCDLPAFQQI 467

Qy        434 ---CSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490
                   |          || |||::: : | :   |    ||||||| | ||||||:|:
Db        468 IPVCQ-------PKNIVKLLNEVFFKLDRIVVLRG---VYKVETVSDSYMTVSGIPDYTS 517
```

```
Qy        491 HHARSICHLALDMMEIAGQVQ--VDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVN 548
Db        518 EHAENMCHVALGMMWEARSVMDPVNKTPFLLRIGLHSGTIIAGVVGTKMPRYCLFGETVT 577

Qy        549 LTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWFLSR-- 606
Db        578 LASQMESLGVAGKIQCSSWTY----SKAMETGRFEFSPRGRINVKGRGD-VETYFLMRSL 632

Qy        607 KNTGTEETKQDDD 619
Db        633 KKSIWEITDHERD 645


RESULT 13
O76340
ID   O76340      PRELIMINARY;      PRT;   940 AA.
AC   O76340;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Soluble guanylyl cyclase beta-3.
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingiodea; Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99107914; PubMed=9891024;
RA   Nighorn A., Byrnes K.A., Morton D.B.;
RT   "Identification and characterization of a novel beta subunit of
RT   soluble guanylyl cyclase that is active in the absence of a second
RT   subunit and Is relatively insensitive to nitric oxide.";
RL   J. Biol. Chem. 274:2525-2531(1999).
DR   EMBL; AF064514; AAD09836.1; -.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR001525; C5_DNA_meth.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00095; C5_MTASE_2; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ   SEQUENCE   940 AA;  106393 MW;  08956886667BD0F6 CRC64;

  Query Match           29.6%;  Score 955;  DB 5;  Length 940;
  Best Local Similarity   35.1%;  Pred. No. 2e-63;
  Matches  221;  Conservative  123;  Mismatches  234;  Indels   52;  Gaps   13;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
Db          1 MYGLLLENMAEYIRQTYGEERWEDIRRQAGV-EQPSFSVHQVYPENLITRLAKKAQEVLG 59

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
Db         60 ITEREFMDQMGVYFVGFVSQYGYDRVLSVLGRHMRDFLNGLDNLHEYLKFSYPRMRAPSF 119

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQR---------- 170
Db        120 IC-ENETRQGLTLHYRSKRRGFVYYAMGQIREVARHFYHKEMRIELLREELLFDTVHVTF 178

Qy        171 NEECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVT 230
Db        179 QLTFDNRAFTLASLAMTREE------------KHLPISASVLFEIFPFCIVFGSDMVVR 225

Qy        231 QCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDV--- 287
Db        226 SIGNSLMVILPDLV--GKKITNWFDLVRPLIAFKFQTILNRTNNIFELVTVEAVMHEKAP 283

Qy        288 ----EKLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDI 343
Db        284 DKRNELLRLSDETDGTTEKNLRLKGQMIYMDNWRMMMYLGTPVMPDLSALVSTGLYINDL 343

Qy        344 PLHDATRDLVLLGEQFREEYKLTQELE-ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSV 402
Db        344 SMHDFSRDLMLAGTQQSVELKLALDQEQQKSKKLEESMRKLDEEMKRTDELLYQMIPKQV 403

Qy        403 ANELRH-KRPV-PAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTL 460
Db        404 ADRLRNGENPIDTCEMFDSVSILFSDVVTFTEICSRIT----PMEVVSMLNAMYSIFDTL 459

Qy        461 TDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQ--VDGESVQ 518
Db        460 TERNR---VYKVETIGDAYMVVSGAPEKEDNHAEKVCDMALDMVDAITDLKDPSTGSHLS 516

Qy        519 ITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENS 578
Db        517 IRVGVHSGAVVAGIVGLKMPRYCLFGDSVNTASRMESTSEAMRIHISQTTQELL------ 570

Qy        579 DPQFHLEHRGPVSMKGKKEPMQVWFLSRKN 608
Db        571 SPSYMVTERGEIQVKGKGAMKTYWLEGRES 600


RESULT 14
O02298
ID   O02298      PRELIMINARY;      PRT;   751 AA.
AC   O02298;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   T04D3.4 protein.
GN   T04D3.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kershaw J.K.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81114; CAB03288.1; -.
DR   HSSP; P19687; 1AWN.
DR   InterPro; IPR001054; G_cyclase.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   Lyase.
SQ   SEQUENCE   751 AA;  85380 MW;  BC864A01AF588CE9 CRC64;

  Query Match           29.0%;  Score 938;  DB 5;  Length 751;
  Best Local Similarity   35.0%;  Pred. No. 2.8e-62;
  Matches  220;  Conservative  125;  Mismatches  236;  Indels   48;  Gaps   15;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
Db          1 MFGWIHESFRQLVTRKYGKDIWEKIVHMSKFELGTESEIAHYYNDDETLRLVNSMANVIG 60

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALH---DHLATIY-PGMR 116
Db         61 IPIEEIWEAYGGFLIQFTMETGWDELLRAMAPDLEGFLDSLDSLHYFIDHV--VYKTKLR 118

Qy        117 APSFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEECD- 175
```

```
Db        119 GPSFRC-DVQADGTLLLHYYSKRSGLYPIVKGVVREVARRIYDTEVVMKVQERKQEHLDA 177

Qy        176 ----HTQFLIEEKE----------SKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHI 221
                  |  |:| : |          | : |   ||       |  |  ||   |  |||:||
Db        178 FVTEHVVFVITQIENANSTQPKSISSKADSQIDL----STGIYE--ISSSDFSLAFPYHI 231

Qy        222 IFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSK 281
               || || |   || | : |       :  :  || | :  |:  |  : |::|| | |
Db        232 CFDPDLFVEHFGNFIKKTFPNAMRQETRVTDLLELVHPEVPFSYESIKYYKNSLFVFRLK 291

Qy        282 EGLLDVEKLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLS 341
               || |:          |   : ||| |::: |   ||::|| :|  : :|  | |:||
Db        292 -GLGDIV------HNANDEAKTVLLKGSMVFIDEGKYILYMCSVNVTTVRELIERNLHLS 344

Qy        342 DIPLHDATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPS 401
              |:  || |||:::| :    : :| : ||  | :|:   : || ||:||| ||  ::| |
Db        345 DMQRHDGTRDVIMLNQSRMSQVELNRTLEETTKKLKKMAQELEIEKQKTDELLCELMPAS 404

Qy        402 VANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLT 461
              ||: ||  : : || : : |:||: || |   |:     :    :| |||||| ||| |
Db        405 VADSLRSGKAMDAKEFADCTLLFTDIVTFTNICAMCTPYD----VVTLLNDLYLRFDRLV 460

Qy        462 DSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV--QVDGESVQI 519
                   |||||:|| || | |:|| | :||  : :::: |:  :  |   :  : ::|
Db        461 GLHD---AYKVETIGDAYMIVGGVPERCENHAERVLNISIGMLMESKLVLSPITHKPIKI 517

Qy        520 TIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSD 579
               :|:| | || ||:| :|||||||:|||: :: |: | : ||:|||      ::   ::
Db        518 RLGVHCGPVVAGVVGIKMPRYCLFGDTVNVANKMESNGIQCKIHVSE---TGKLNGLKAN 574

Qy        580 PQFHLEHRGPVSMKGKKEPMQVWFLSRKN 608
              | :    ||   ::| | | :|| | :
Db        575 PSYVFIDRGNTEIRG-KGMMYTYFLERND 602
```

```
RESULT 15
Q9VFC5
ID   Q9VFC5      PRELIMINARY;      PRT;   690 AA.
AC   Q9VFC5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   CG4154 protein.
GN   CG4154.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003707; AAF55135.1; -.
DR   HSSP; Q02846; 1AWL.
DR   FlyBase; FBgn0038295; CG4154.
DR   InterPro; IPR001525; C5_DNA_meth.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS00095; C5_MTASE_2; UNKNOWN_1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   690 AA;  78736 MW;  4B736D7413E168AA CRC64;

  Query Match           28.8%;  Score 929.5;  DB 5;  Length 690;
  Best Local Similarity   34.8%;  Pred. No. 1.1e-61;
  Matches  218;  Conservative  129;  Mismatches  231;  Indels   49;  Gaps   15;

Qy          1 MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLN 60
              ||| :   |   :   || | ||||:::| :|    | |  :| ::    |   | :||
Db          1 MYGLLLENLSEYIKSVYGEEKWEDIRRQAGIDSP-SFSVHQVYPENLLQKLAKKAQQVLG 59

Qy         61 LNAGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSF 120
              ::  : :   |  |  |  : ||| :| ||| ::|:||  || ||::|   || ||||||
Db         60 VSERDFMDQMGVYFVGFVGQYGYDRVLSVLGRHMRDFLNGLDNLHEYLKFSYPRMRAPSF 119

Qy        121 RCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQR---------- 170
               | :  | :|| ||| |:| |     :| |: ||:  :  |: ::::::
Db        120 ICENETK-QGLTLHYRSKRRGFVYYTMGQIREVARYFYHKEMHIELVREEILFDTVHVTF 178

Qy        171 NEECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFDRDLVVT 230
                 |:  | :       |:           :   || :   : ||| ::|  |:||
Db        179 QLTFDNRAFTLASLAMTREE-----------KHLPISAHVLFEIFPFCMVFGADMVVR 225

Qy        231 QCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKL 290
               ||::  :||:|      : : | |||| |   |  ||:  | :| | : : : :   :
Db        226 SIGNSLMVILPELL--GKKITAWFDLVRPLIAFKFQTILNRTNNIFELVTVDPVTERFDV 283

Qy        291 ECEDEL---TGTE-ISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346
              : || |    |:|    |||||||:|:     |:|| :| : :|  |   |||::|: :|
Db        284 QNEDLLQHEDGSEPEKSLRLKGQMVYMENWRMIMFLGTPVMPDLTSLITTGLYINDLSMH 343

Qy        347 DATRDLVLLGEQFREEYKLTQELE-ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANE 405
              | :|||:| | |   | ||  : |   : :|: ::| |::| ::|| ||| ::|  ||:
Db        344 DFSRDLMLAGTQQSVELKLALDQEQQKSKKLEESMRLLDEEMRRTDELLYQMIPKQVADR 403

Qy        406 LRH-KRPV-PAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDS 463
              ||  : |:   : :|:|:|||| || |   ||:       |::|::|| :|: || ||:
Db        404 LRRGENPIDTCEMFDSVSILFSDIVTFTEICSRIT----PMEVVSMLNAMYSIFDKLTER 459

Qy        464 RKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQ--VDGESVQITI 521
                    |||||:|| || |:| |:   :||  :| :||||::    ::    |: ::| :
Db        460 NS---VYKVETIGDAYMVVAGAPDKDANHAERVCDMALDMVDAITDLKDPSTGQHLRIRV 516

Qy        522 GIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQ 581
              |:|:| || |::| :||||||||:|||  || |:|    |:::|| | : |:      |
Db        517 GVHSGAVVAGIVGLKMPRYCLFGDTVNTASRMESTSIAMKVHISEST-KVLIG-----PN 570
```

```
Qy        582 FHLEHRGPVSMKGKKEPMQVWFLSRKN 608
                : :  || : :|||      |   |:|
Db        571 YKIIERGEIDVKGKGTMGTYWLEEREN 597
```

Search completed: June 27, 2003, 13:04:43
Job time : 33.2101 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:32 ; Search time 12.7678 Seconds
(without alignments)
1426.467 Million cell updates/sec

Title: US-09-762-767A-4
Perfect score: 3231
Sequence: 1 MYGFVNHALELLVIRNYGPE.....QVWFLSRKNTGTEETKQDDD 619

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	404	12.5	1075	1	US-07-623-033-2	Sequence 2, Appli
2	344.5	10.7	1253	4	US-08-864-785-2	Sequence 2, Appli
3	309.5	9.6	1134	3	US-08-726-214-2	Sequence 2, Appli
4	309	9.6	533	3	US-08-726-214-18	Sequence 18, Appl
5	308.5	9.5	1248	3	US-08-726-214-16	Sequence 16, Appl
6	291.5	9.0	1294	4	US-09-473-717-2	Sequence 2, Appli
7	291.5	9.0	1305	4	US-08-864-785-3	Sequence 3, Appli
8	291.5	9.0	1353	3	US-08-894-173-2	Sequence 2, Appli
9	291.5	9.0	1353	4	US-09-398-193-2	Sequence 2, Appli
10	291.5	9.0	1353	4	US-09-473-717-3	Sequence 3, Appli
11	288	8.9	1098	3	US-08-726-214-10	Sequence 10, Appl
12	286.5	8.9	1144	3	US-08-726-214-6	Sequence 6, Appli
13	283.5	8.8	1353	4	US-09-398-193-99	Sequence 99, Appl
14	282.5	8.7	1099	3	US-08-726-214-14	Sequence 14, Appl
15	282.5	8.7	1168	4	US-09-474-076-2	Sequence 2, Appli
16	281	8.7	1064	3	US-08-726-214-8	Sequence 8, Appli
17	280	8.7	1180	3	US-08-726-214-12	Sequence 12, Appl
18	278.5	8.6	604	4	US-09-008-097-4	Sequence 4, Appli
19	278.5	8.6	1167	4	US-09-008-097-6	Sequence 6, Appli
20	278	8.6	1261	4	US-09-473-716-2	Sequence 2, Appli
21	272.5	8.4	1165	1	US-08-240-357-2	Sequence 2, Appli
22	271.5	8.4	1090	3	US-08-307-896-3	Sequence 3, Appli
23	271.5	8.4	1090	3	US-08-726-214-4	Sequence 4, Appli
24	271.5	8.4	1090	5	PCT-US95-11808-3	Sequence 3, Appli
25	269.5	8.3	1077	4	US-09-412-210-1	Sequence 1, Appli
26	252.5	7.8	429	3	US-08-307-896-4	Sequence 4, Appli
27	252.5	7.8	429	5	PCT-US95-11808-4	Sequence 4, Appli
28	175.5	5.4	670	2	US-08-997-080-178	Sequence 178, App
29	175.5	5.4	670	2	US-08-997-362-178	Sequence 178, App
30	175.5	5.4	670	4	US-09-095-855-178	Sequence 178, App
31	175.5	5.4	670	4	US-09-324-542-178	Sequence 178, App
32	175.5	5.4	670	4	US-09-205-426-178	Sequence 178, App
33	175.5	5.4	722	2	US-08-997-080-174	Sequence 174, App
34	175.5	5.4	722	2	US-08-997-362-174	Sequence 174, App
35	175.5	5.4	722	4	US-09-095-855-174	Sequence 174, App
36	175.5	5.4	722	4	US-09-324-542-174	Sequence 174, App
37	175.5	5.4	722	4	US-09-205-426-174	Sequence 174, App
38	173	5.4	419	2	US-08-997-080-125	Sequence 125, App
39	173	5.4	419	2	US-08-997-362-125	Sequence 125, App
40	173	5.4	419	4	US-09-095-855-125	Sequence 125, App
41	173	5.4	419	4	US-09-324-542-125	Sequence 125, App
42	173	5.4	419	4	US-09-205-426-125	Sequence 125, App
43	172.5	5.3	112	4	US-08-857-076-75	Sequence 75, Appl
44	149	4.6	107	4	US-08-857-076-76	Sequence 76, Appl
45	137.5	4.3	88	3	US-08-894-173-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-07-623-033-2
; Sequence 2, Application US/07623033
; Patent No. 5237051
; GENERAL INFORMATION:
; APPLICANT: Garbers, David L.
; APPLICANT: Schulz, Stephanie
; TITLE OF INVENTION: CLONING THE ENTEROTOXIN RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGMUS & CHESTNUT
; STREET: 100 South Wacker Drive - Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/623,033
; FILING DATE: 19901206
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: VU9018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 456-8000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1075 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-623-033-2

Query Match 12.5%; Score 404; DB 1; Length 1075;
Best Local Similarity 37.8%; Pred. No. 4.8e-31;
Matches 102; Conservative 40; Mismatches 98; Indels 30; Gaps 8;

```
Qy      354 LLGEQFREEYKLTQELEILTDRLQLTLRALED----------EKKKTDTLLYSVLPPSV 402
            |  :|  | |     :: |  ||||  | ||              |: : | | : :||  |
Db      753 LFHDQKNESY-----MDTLIRRLQLYSRNLEHLVEERTQLYKAERDRADHLNFMLLPRLV 807

Qy      403 ANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTD 462
                |: |  |  : |: ||| || ||||   | |:::    |::|::|||:|  || : |
Db      808 VKSLKEKGIVEPELYEEVTIYFSDIVGFTTIC-KYST---PMEVVDMLNDIYKSFDQIVD 863
```

```
Qy        463 SRKNPFVYKVETVGDKYMTVSGLP-EPCIHHARSICHLALDMMEIAGQVQVD---GESVQ 518
                  ||||||:|| |:  ||||      ||  |  :|||::   |  :::   |  |
Db        864 HHD---VYKVETIGDAYVVASGLPMRNGNRHAVDISKMALDILSFMGTFELEHLPGLPVW 920

Qy        519 ITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENS 578
              | ||:|:|    ||:| :|||||||||:|||  || |:||   :|::|  |   |     :
Db        921 IRIGVHSGPCAAGVVGIKMPRYCLFGDTVNTASRMESTGLPLRIHMSSSTIAIL---RRT 977

Qy        579 DPQFHLEHRGPVSMKGKKEPMQVWFLSRKN 608
              | ||  | ||   :||:      |    |:
Db        978 DCQFLYEVRGETYLKGRGTETTYWLTGMKD 1007


RESULT 2
US-08-864-785-2
; Sequence 2, Application US/08864785A
; Patent No. 6329566
; GENERAL INFORMATION:
;  APPLICANT: Kaplan, Joshua M.
;  APPLICANT: Oppenheimer, Allison J.
;  APPLICANT: Hart, Anne C.
;  TITLE OF INVENTION: METHODS FOR THE DETECTION, TREATMENT,
;  TITLE OF INVENTION: AND PREVENTION OF NEURODEGENERATION
;  FILE REFERENCE: 00786/353001
;  CURRENT APPLICATION NUMBER: US/08/864,785A
;  CURRENT FILING DATE: 1997-05-29
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1253
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-08-864-785-2

  Query Match             10.7%;  Score 344.5;  DB 4;  Length 1253;
  Best Local Similarity   32.5%;  Pred. No. 5.6e-25;
  Matches   93;  Conservative   45;  Mismatches   85;  Indels   63;  Gaps    8;

Qy        378 LTLRALEDEKKKTDTLLYSVLPPSVANEL---------------------------- 406
              |  : || | :  | :: ||:|  ||:||
Db        239 LARKDLELETQFKDHMIQSVMPKKVADELLKDASELRRPSASNDSNCRTSNATQVDQPLA 298

Qy        407 ------RHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTL 460
                  |  ||       ||:|||: | ||    :| :| : | ::||||||: |||||
Db        299 KMVPEYRKFRPFTMNLMTNVSILFADIAGF----TKMSSNKSADELVNLLNDLFGRFDTL 354

Qy        461 TDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQVQVD-GESVQI 519
              |    : |: |:|| |  |:| ||||  ||     : |||:    |  :| |: | :
Db        355 CRLRG---LEKISTLGDCYYCVAGCPEPCDDHACRTVEMGLDMIVAIRQFDIDRGQEVNM 411

Qy        520 TIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMS----P 575
               :|||||:|: |::| :  :: :| | | | :  |::|  |:::||| | : |
Db        412 RVGIHTGKVMCGMVGTKRFKFDVFSNDVTLANEMESSGVAGRVHVSEATAKLLKGLYEIE 471

Qy        576 ENSDPQFHLEHRGPVSM-------KGKKEPMQVWFL-SRKNTGTEE 613
              |  |      : ||: |       : | | |: :|:  | | | ||
Db        472 EGPD------YDGPLRMQVQGTERRVKPESMKTFFIKGRINDGVEE 511


RESULT 3
US-08-726-214-2
; Sequence 2, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1134 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-2

  Query Match             9.6%;  Score 309.5;  DB 3;  Length 1134;
  Best Local Similarity   33.1%;  Pred. No. 1.5e-21;
  Matches   80;  Conservative   49;  Mismatches   80;  Indels   33;  Gaps    8;

Qy        354 LLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELR------ 407
              :| |: : :  | |    : |||:     |||| :| : || |:|| :|| |::
Db        235 ILAERAQRKAFL-QARNCIEDRLR-----LEDENEKQERLLMSLLPRNVAMEMKEDFLKP 288

Qy        408 -----HKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTD 462
                  ||  :  :|:|||:|||: ||||    |:       | ::| |||:|: :|| |
Db        289 PERIFHK--IYIQRHDNVSILFADIVGFTGLASQCT----AQELVKLLNELFGKFDELAT 342

Qy        463 SRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMME----IAGQVQVDGESVQ 518
                    ::: :|| |  |||| :|   ||     : |||::    :|   :||   :
Db        343 EN---HCRRIKILGDCYYCVSGLTQPKTDHAHCCVEMGLDMIDTITSVAEATEVD---LN 396

Qy        519 ITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENS 578
              : :|:||| |: ||:| |  :| :: | | | :  |  |  ||::::: |  ||
Db        397 MRVGLHTGRVLCGVLGLRKWQYDVWSNDVTLANVMEAAGLPGKVHITKTTLACLNGDYEV 456

Qy        579 DP 580
              :|
Db        457 EP 458


RESULT 4
US-08-726-214-18
; Sequence 18, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  533 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-18

  Query Match             9.6%;  Score 309;  DB 3;  Length 533;
  Best Local Similarity   33.6%;  Pred. No. 4.6e-22;
  Matches   77;  Conservative   45;  Mismatches   75;  Indels   32;  Gaps    7;

Qy        367 QELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELR----------HKRPVPAK 415
              |    : |||:     |||| :| : || |:|| :|| |::           ||  :  :
Db         12 QARNCIEDRLR-----LEDENEKQERLLMSLLPRNVAMEMKEDFLKPPERIFHK--IYIQ 64

Qy        416 RYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETV 475
              |:|||:|||: |||    |:      | ::| |||:|: :|| |          ::: :
Db         65 RHDNVSILFADIVGFTGLASQCT----AQELVKLLNELFGKFDELATEN---HCRRIKIL 117

Qy        476 GDKYMTVSGLPEPCIHHARSICHLALDMME----IAGQVQVDGESVQITIGIHTGEVVTG 531
              || |  |||| :|   ||      : |||::    :|   :||   : : :|:|||  |: |
Db        118 GDCYYCVSGLTQPKTDHAHCCVEMGLDMIDTITSVAEATEVD---LNMRVGLHTGRVLCG 174

Qy        532 VIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDP 580
              |:| |  :| :: | | | :  |  |  ||:::::  |  ||      :|
Db        175 VLGLRKWQYDVWSNDVTLANVMEAAGLPGKVHITKTTLACLNGDYEVEP 223


RESULT 5
US-08-726-214-16
; Sequence 16, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1248 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-16

  Query Match             9.5%;  Score 308.5;  DB 3;  Length 1248;
  Best Local Similarity   25.3%;  Pred. No. 2.2e-21;
  Matches  136;  Conservative   86;  Mismatches  209;  Indels  107;  Gaps   22;

Qy        153 VAQQIH---GTEIDMKVIQQRNEECDHT--------------QFLIEEKESKEEDFYED 194
              :|| :|   | |       : |:  :||               | :  |:|  |  |
Db        651 LAQALHVQSGPE-------EINKRIEHTIDLRSGDKLRREHIKPFSLMFKDSSLEHKYSQ 703

Qy        195 L--DRFEENGT---------------QESRISPYTFCKAFPFHIIFDRDLVVTQCGNAI 236
              :  : |: |                    ||: | |    |   |:    ||:
Db        704 MRDEVFKSNLVCAFIVLLFITAIQSLLPSSRLMPMTI--QFSILIMLHSALVLITTAED- 760

Qy        237 YRVLPQLQPGNCSLLSVFSLVR-----PHIDISFHGILSHINTVFVLRSKEGLLDVEKLE 291
              |: || :    |  ::   | |       | |:| | :  || ::    |   : :: |
Db        761 YKCLPLILRKTCCWINETYLARNVIIFASILINFLGAV--INILWCDFDKS--IPLKNLT 816

Qy        292 CEDELTGTEI----------SCLRLKGQMIYLPEADSILFLCSPSVM-----NLDDLTRR 336
                    |:|            | :    ::|   :|:| |    :|      | :
Db        817 FNSSAVFTDICSYPEYFVFTGVLAMVTCAVFL-RLNSVLKLAVLLIMIAIYALLTETIYA 875

Qy        337 GLYLSDIPLHDATRD--------LVLLGEQFREEYKLTQELEILTDRLQLTLRA------ 382
              ||:||   |: :  |       |:|:       :   |:||  | ||    |
Db        876 GLFLSYDNLNHSGEDFLGTKEASLLLMAMFLLAVFYHGQQLE-YTARLDFLWRVQAKEEI 934

Qy        383 --LEDEKKKTDTLLYSVLPPSVANEL----RHKRPVPAKRYDNVTILFSGIVGFNAFCSK 436
                ::| ::  : :| ::||  ||       |    : :: || | ::|: | ||  | |:
Db        935 NEMKDLREHNENMLRNILPGHVARHFLEKDRDNEELYSQSYDAVGVMFASIPGFADFYSQ 994

Qy        437 HASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGL-PE--PCIHHA 493
                  :: : |||::   || |    :  : |::|:|  || |||| ||   |
Db        995 TEMNNQGVECLRLLNEIIADFDELLGEDRFQDIEKIKTIGSTYMAVSGLSPEKQQCEDKW 1054

Qy        494 RSICHLALDMMEIAGQVQVDGE----SVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549
              :| ||   : :   :|   :    : :: |||  | || |||| : |:| ::| ||||
Db       1055 GHLCALADFSLALTESIQEINKHSFNNFELRIGISHGSVVAGVIGAKKPQYDIWGKTVNL 1114

Qy        550 TSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKE---PMQVWFL 604
               || ::||  |:| | | ||  |      |  |  ::|| : :||  |    :: :||
Db       1115 ASRMDSTGVSGRIQVPEETYLIL-----KDQGFAFDYRGEIYVKGISEQEGKIKTYFL 1167


RESULT 6
US-09-473-717-2
; Sequence 2, Application US/09473717
; Patent No. 6372475
; GENERAL INFORMATION:
;  APPLICANT:  Storm, Daniel R.
;  APPLICANT:  Hacker, Beth
;  APPLICANT:  Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  APPLICANT:  University of Washington
;  TITLE OF INVENTION:  CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE:  44481-5029-01-US
```

```
;  CURRENT APPLICATION NUMBER: US/09/473,717
;  CURRENT FILING DATE:  1999-12-29
;  PRIOR APPLICATION NUMBER:  PCT/US98/13541
;  PRIOR FILING DATE:  1998-07-01
;  PRIOR APPLICATION NUMBER:  60/098,559
;  PRIOR FILING DATE:  1997-07-01
;  PRIOR APPLICATION NUMBER:  08/886,440
;  PRIOR FILING DATE:  1997-07-01
;  NUMBER OF SEQ ID NOS:  3
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1294
;   TYPE: PRT
;   ORGANISM: human type IX adenylyl cyclase
US-09-473-717-2

  Query Match             9.0%;  Score 291.5;  DB 4;  Length 1294;
  Best Local Similarity   31.2%;  Pred. No. 1.2e-19;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL-------------RHK------------------- 409
              || ||   : :::||:|  :|::|              ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEETK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
              |: |:|| |  |:| |||   ||     : | |::   |  |   | | : :|:||| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 7
US-08-864-785-3
; Sequence 3, Application US/08864785A
; Patent No. 6329566
; GENERAL INFORMATION:
;  APPLICANT: Kaplan, Joshua M.
;  APPLICANT: Oppenheimer, Allison J.
;  APPLICANT: Hart, Anne C.
;  TITLE OF INVENTION: METHODS FOR THE DETECTION, TREATMENT,
;  TITLE OF INVENTION: AND PREVENTION OF NEURODEGENERATION
;  FILE REFERENCE: 00786/353001
;  CURRENT APPLICATION NUMBER: US/08/864,785A
;  CURRENT FILING DATE: 1997-05-29
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1305
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-08-864-785-3

  Query Match             9.0%;  Score 291.5;  DB 4;  Length 1305;
  Best Local Similarity   31.2%;  Pred. No. 1.2e-19;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL-------------RHK------------------- 409
              || ||   : :::||:|  :|::|              ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEQTK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
              |: |:|| |  |:| |||   ||     : | |::   |  |   | | : :|:||| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 8
US-08-894-173-2
; Sequence 2, Application US/08894173A
; Patent No. 6090612
; GENERAL INFORMATION:
;  APPLICANT: Medical Research Council
;  TITLE OF INVENTION: Adenylate cyclase and uses therefor
;  FILE REFERENCE: P14716C
;  CURRENT APPLICATION NUMBER: US/08/894,173A
;  CURRENT FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: Mouse
;   FEATURE:
;   OTHER INFORMATION: HYPOTHETICAL : NO
US-08-894-173-2

  Query Match             9.0%;  Score 291.5;  DB 3;  Length 1353;
  Best Local Similarity   31.2%;  Pred. No. 1.3e-19;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL-------------RHK------------------- 409
              || ||   : :::||:|  :|::|              ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEQTK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
              |: |:|| |  |:| |||   ||     : | |::   |  |   | | : :|:||| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 9
US-09-398-193-2
; Sequence 2, Application US/09398193
; Patent No. 6197581
; GENERAL INFORMATION:
;  APPLICANT: Medical Research Council
;  TITLE OF INVENTION: Adenylate cyclase and uses therefor
;  FILE REFERENCE: P24360-
;  CURRENT APPLICATION NUMBER: US/09/398,193
;  CURRENT FILING DATE: 1999-09-17
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: Mouse
;   FEATURE:
;   OTHER INFORMATION: HYPOTHETICAL : NO
US-09-398-193-2

  Query Match             9.0%;  Score 291.5;  DB 4;  Length 1353;
  Best Local Similarity   31.2%;  Pred. No. 1.3e-19;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL-------------RHK------------------- 409
              || ||   : :::||:|  :|::|              ||
```

```
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEQTK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
              |: |:|| |  |:| |||   ||     : | |::   |  |   | | : :|:||| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 10
US-09-473-717-3
; Sequence 3, Application US/09473717
; Patent No. 6372475
; GENERAL INFORMATION:
;  APPLICANT:  Storm, Daniel R.
;  APPLICANT:  Hacker, Beth
;  APPLICANT:  Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  APPLICANT:  University of Washington
;  TITLE OF INVENTION:  CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE:  44481-5029-01-US
;  CURRENT APPLICATION NUMBER: US/09/473,717
;  CURRENT FILING DATE:  1999-12-29
;  PRIOR APPLICATION NUMBER:  PCT/US98/13541
;  PRIOR FILING DATE:  1998-07-01
;  PRIOR APPLICATION NUMBER:  60/098,559
;  PRIOR FILING DATE:  1997-07-01
;  PRIOR APPLICATION NUMBER:  08/886,440
;  PRIOR FILING DATE:  1997-07-01
;  NUMBER OF SEQ ID NOS:  3
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: murine type IX adenylyl cyclase
US-09-473-717-3

  Query Match             9.0%;  Score 291.5;  DB 4;  Length 1353;
  Best Local Similarity   31.2%;  Pred. No. 1.3e-19;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL------------RHK------------------- 409
              || ||   : :::||:|  :|::|             ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEQTK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
              |: |:|| |  |:| |||   ||     : | |::   |  |   | | : :|:||| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 11
US-08-726-214-10
; Sequence 10, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1098 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-10

  Query Match             8.9%;  Score 288;  DB 3;  Length 1098;
  Best Local Similarity   23.5%;  Pred. No. 2e-19;
  Matches  103;  Conservative   84;  Mismatches  153;  Indels   98;  Gaps   16;

Qy        248 CSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLEC------EDELTGTEI 301
              | ::  |   || : : :          : || :  |:: :  : |      :| :    :
Db         90 CLVMLAFHAARPPLQVVY---------LAVLAAAVGVILIMAVLCNRAAFHQDHM---GL 137

Qy        302 SC---------LRLKGQMIYLPEADS------ILFLCSPSVMNLDDLTRRGLYLSD---- 342
              :|         ::: | ::  | : |      : |:   ::  |  :  |   ||
Db        138 ACYALIAVVLAVQVVGLLLPQPRSASEGIWWTVFFIY--TIYTLLPVRMRAAVLSGVLLS 195

Qy        343 -----IPLHDATRDLVLLGE-----------------QFREEYKLTQELEILTDRLQLT 379
                  | ||   :|  || :                 :  |    |  :   : :|
Db        196 ALHLAISLHTNAQDQFLLKQLVSNVLIFSCTNIVGVCTHYPAEVSQRQAFQETRECIQAR 255

Qy        380 LRALEDEKKKTDTLLYSVLPPSVANELR-----------HKRPVPAKRYDNVTILFSGI 427
              | : : | :: : || ||||  || |::            ||  :  :::|||:|||: |
Db        256 LHS-QRENQQQERLLLSVLPRHVAMEMKADINAKQEDMMFHK--IYIQKHDNVSILFADI 312

Qy        428 VGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPE 487
              || :  |:       | ::|  ||:|: ||| |          ::: :|| |  ||||||
Db        313 EGFTSLASQCT----AQELVMTLNELFARFDKLAAEN---HCLRIKILGDCYYCVSGLPE 365

Qy        488 PCIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNT 546
                    ||      : :||:|   | :| | :| : :|||:| |  ||:| |  :: :: |
Db        366 ARADHAHCCVEMGMDMIEAISSVREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSND 425

Qy        547 VNLTSRTETTGEKGKINVSEYTYRCLMSPENSDP-------QFHLEHRGPV------SMK 593
              | | :  |  |: |:|:::: |   |       :|        :  ||.         : |
Db        426 VTLANHMEAGGKAGRIHITKATLNYLNGDYEVEPGCGGERNAYLKEHSIETFLILRCTQK 485

Qy        594 GKKEPMQVWFLSRKNTGT 611
              |:|   :  ::|: | :
Db        486 RKEEKAMIAKMNRQRTNS 503
```

```
RESULT 12
US-08-726-214-6
; Sequence 6, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1144 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-6

  Query Match             8.9%;  Score 286.5;  DB 3;  Length 1144;
  Best Local Similarity   21.8%;  Pred. No. 3.1e-19;
  Matches  119;  Conservative  104;  Mismatches  201;  Indels  121;  Gaps   19;

Qy        165 KVIQQRNEECDHTQFLIEEKESKEEDFYEDLDRFEENGTQESRISPYTFCKAFPFHIIFD 224
              :|:::||      :|:  | |::       ::: :::|   |      || |    |: |
Db        599 QVVKKRNTFLLTMRFMDPEMETR-----YSVEKEKQSGAAFSCSCVVLFCTAM-VEILID 652

Qy        225 ---------------RDLVVTQCGNAIYRVLPQLQP----------------GNCSLLS 252
                           |::| |  |   : |: |                     ::|:
Db        653 PWLMTNYVTFVVGEVLLLILTICSMA--AIFPRAFPKKLVAFSSWIDRTRWARNTWAMLA 710

Qy        253 VFSLVRPH-IDI-----------------------------SFHGILSHINTVFVLRSK 281
              :| ||  : :|:                          ::  :|| | |: :::
Db        711 IFILVMANVVDMLSCLQYYMGPYNVTTGIELDGGCMENPKYYNYVAVLSLIATIMLVQ-- 768

Qy        282 EGLLDVEKLECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVM----NLDDLTRRG 337
              :  : ||      :||  :: : |        | |   |    | |     :  |: |
Db        769 --VSHMVKLTLMLLVTGA-VTAINLYAWCPVFDEYDHKRFQEKDSPMVALEKMQVLSTPG 825

Qy        338 LYLSD--IPLHDATRDL-VLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKK----- 389
              |  :|  :|| :   : |::       |  :: :| |   | |    : |:|::
Db        826 LNGTDSRLPLVPSKYSMTVMMFVMMLSFYYFSRHVEKLARTLFLWKIEVHDQKERVYEMR 885

Qy        390 --TDTLLYSVLPPSVANEL----RHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGA 443
                : |: ::||  ||       :    : :: || : ::|: :  |  | :: :   |
Db        886 RWNEALVTNMLPEHVARHFLGSKKRDEELYSQSYDEIGVMFASLPNFADFYTEESINNGG 945

Qy        444 MKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIH------------ 491
              :: :  ||:: : ||:| |: |   : |::|:|  ||  ||:  | ::
Db        946 IECLRFLNEIISDFDSLLDNPKFRVITKIKTIGSTYMAASGV-TPDVNTNGFTSSSKEEK 1004

Qy        492 -------HARSICHLALDMMEIAGQVQVDG-ESVQITIGIHTGEVVTGVIGQRMPRYCLF 543
                    |   :   || | :    :       :  : ||:: | |: |||| | | | ::
Db       1005 SDKERWQHLADLADFALAMKDTLTNINNQSFNNFMLRIGMNKGGVLAGVIGARKPHYDIW 1064

Qy        544 GNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPMQVWF 603
              |||||: || |:||  | | | | |   |      :  |    |||: :||| | :  :|
Db       1065 GNTVNVASRMESTGVMGNIQVVEETQVIL-----REYGFRFVRRGPIFVKGKGE-LLTFF 1118

Qy        604 LSRKN 608
              |  ::
Db       1119 LKGRD 1123


RESULT 13
US-09-398-193-99
; Sequence 99, Application US/09398193
; Patent No. 6197581
; GENERAL INFORMATION:
;  APPLICANT: Medical Research Council
;  TITLE OF INVENTION: Adenylate cyclase and uses therefor
;  FILE REFERENCE: P24360-
;  CURRENT APPLICATION NUMBER: US/09/398,193
;  CURRENT FILING DATE: 1999-09-17
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 99
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: Human
US-09-398-193-99

  Query Match             8.8%;  Score 283.5;  DB 4;  Length 1353;
  Best Local Similarity   30.8%;  Pred. No. 8.2e-19;
  Matches   73;  Conservative   44;  Mismatches   73;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL------------RHK------------------- 409
              || ||   : :::||:|  :|::|              ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEETK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
              |: |:|| |  |:| |||   ||     : | |::   |  |   | | : :|:||  |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTRTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 14
US-08-726-214-14
; Sequence 14, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,214
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,498
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1099 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-726-214-14

  Query Match             8.7%;  Score 282.5;  DB 3;  Length 1099;
  Best Local Similarity   29.3%;  Pred. No. 7.2e-19;
  Matches   84;  Conservative   55;  Mismatches   97;  Indels   51;  Gaps    9;

Qy        334 TRRGLYLSDIPLHDATRDLVLLGEQFREEY------KLTQELEILTDRLQLTLRALEDEK 387
              |: || |          ::|||  |   :       :::| | | :     | |  ||
Db        176 TQLGLQL-------LANAVILLGGNFTGAFHKHQLQDASRDLFIYTVKCIQIRRKLRVEK 228

Qy        388 KKTDTLLYSVLPPSVANEL-------------RHKRP------VPAKRYDNVTILFSGIV 428
              :: : || ||||  ::  :             ||  |      :  ||: ||:||:: ||
Db        229 RQQENLLLSVLPAHISMGMKLAIIERLKEGGDRHYMPDNNFHSLYVKRHQNVSILYADIV 288

Qy        429 GFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEP 488
              ||    :: ||     ::| :||:|: :|| :  : :     ::: :|| |  |||||
Db        289 GF----TRLASDCSPKELVVVLNELFGKFDQIAKANE---CMRIKILGDCYYCVSGLPVS 341

Qy        489 CIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTV 547
                 |||:   : ||: |   || :  |  : : :|||:| |: |||| |  :| :: : |
Db        342 LPTHARNCVKMGLDICEAIKQVREATGVDISMRVGIHSGNVLCGVIGLRKWQYDVWSHDV 401

Qy        548 NLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQF-----HLEHRGP 589
              :| :| |  |  |:::::| |   |      |  :     | | | |
Db        402 SLANRMEAAGVPGRVHITEATLNHL------DKAYEVEDGHGEQRDP 442


RESULT 15
US-09-474-076-2
; Sequence 2, Application US/09474076
; Patent No. 6465237
; GENERAL INFORMATION:
;  APPLICANT:  Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  TITLE OF INVENTION:  CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE:  44481-5028-01-US
;  CURRENT APPLICATION NUMBER: US/09/474,076
;  CURRENT FILING DATE:  1999-12-12
;  PRIOR APPLICATION NUMBER:  PCT/US98/13694
;  PRIOR FILING DATE:  1998-07-01
;  PRIOR APPLICATION NUMBER:  60/070,904
;  PRIOR FILING DATE:  1997-07-01
;  PRIOR APPLICATION NUMBER:  08/886,550
;  PRIOR FILING DATE:  1997-07-01
;  NUMBER OF SEQ ID NOS:  2
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1168
;   TYPE: PRT
;   ORGANISM: human type VI adenylyl cyclase
US-09-474-076-2

  Query Match             8.7%;  Score 282.5;  DB 4;  Length 1168;
  Best Local Similarity   26.6%;  Pred. No. 8e-19;
  Matches   94;  Conservative   61;  Mismatches  126;  Indels   73;  Gaps   12;

Qy        296 LTGTEISCLRL-------KGQMIYLPE--ADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346
              |:|  :| | |        :|      :  |: :||||:
Db        262 LSGLGLSTLHLILAWQLNRGDAFLWKQLGANVLLFLCT---------------------- 299

Qy        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406
                  ::: :   :  |    |  :     :|  |  |: | :: : || ||||  || |:
Db        300 ----NVIGICTHYPAEVSQRQAFQETRGYIQARLH-LQHENRQQERLLLSVLPQHVAMEM 354

Qy        407 R------------HKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLY 454
              :            ||  :  :::|||:|||: | || :  |:      | ::|  ||:|:
Db        355 KEDINTKKEDMMFHK--IYIQKHDNVSILFADIEGFTSLASQCT----AQELVMTLNELF 408

Qy        455 TRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVD 513
               ||| |          ::: :|| |  ||||||    ||     : :||:|    | :|
Db        409 ARFDKLAAEN---HCLRIKILGDCYYCVSGLPEARADHAHCCVEMGVDMIEAISLVREVT 465

Qy        514 GESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLM 573
              | :| : :|||:| |  ||:| |  :: :: | | | :  |  |  |:|:::  | : |
Db        466 GVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHMEAGGRAGRIHITRATLQYLN 525

Qy        574 SPENSDP-----------QFHLEH---RGPVSMKGKKEPMQVWFLSRKNTGTEE 613
                  :|           : |:|     |  | | |:|   :  | |    : |
Db        526 GDYEVEPGRGGERNAYLKEQHIETFLILG-ASQKRKEEKAMLAKLQRTRANSME 578


Search completed: June 27, 2003, 13:05:59
Job time : 14.7678 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:05:32 ; Search time 32.6287 Seconds
(without alignments)
2080.112 Million cell updates/sec

Title: US-09-762-767A-4
Perfect score: 3231
Sequence: 1 MYGFVNHALELLVIRNYGPE.....QVWFLSRKNTGTEETKQDDD 619

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	789.5	24.4	690	9	US-10-205-823-162	Sequence 162, App
2	778	24.1	691	9	US-09-952-213D-2	Sequence 2, Appli
3	407.5	12.6	1073	10	US-09-819-249-2	Sequence 2, Appli
4	405.5	12.6	1073	9	US-10-157-031-18	Sequence 18, Appl
5	291.5	9.0	1294	12	US-10-071-223-2	Sequence 2, Appli
6	291.5	9.0	1353	10	US-09-751-100B-2	Sequence 2, Appli
7	291.5	9.0	1353	12	US-10-071-223-3	Sequence 3, Appli
8	289	8.9	1080	9	US-10-282-942-2	Sequence 2, Appli
9	283.5	8.8	1353	10	US-09-751-100B-99	Sequence 99, Appl
10	282.5	8.7	1168	9	US-10-201-000-2	Sequence 2, Appli
11	282.5	8.7	1168	10	US-09-750-240-11	Sequence 11, Appl
12	278.5	8.6	604	10	US-09-750-240-4	Sequence 4, Appli
13	278.5	8.6	1167	10	US-09-750-240-6	Sequence 6, Appli
14	278	8.6	1261	9	US-10-175-158-2	Sequence 2, Appli
15	269.5	8.3	257	9	US-09-989-442-95	Sequence 95, Appl
16	269.5	8.3	302	10	US-09-915-582-50	Sequence 50, Appl
17	269.5	8.3	1077	9	US-10-121-911-1	Sequence 1, Appli
18	265.5	8.2	1167	10	US-09-750-240-13	Sequence 13, Appl
19	264	8.2	241	9	US-09-764-868-794	Sequence 794, App
20	264	8.2	241	9	US-09-989-442-141	Sequence 141, App
21	259.5	8.0	855	10	US-09-925-297-811	Sequence 811, App
22	242	7.5	225	9	US-09-989-442-92	Sequence 92, Appl
23	234.5	7.3	311	10	US-09-925-297-515	Sequence 515, App
24	175.5	5.4	670	9	US-10-051-643-178	Sequence 178, App
25	175.5	5.4	670	9	US-09-880-505-178	Sequence 178, App
26	175.5	5.4	722	9	US-10-051-643-174	Sequence 174, App
27	175.5	5.4	722	9	US-09-880-505-174	Sequence 174, App
28	173	5.4	419	9	US-10-051-643-125	Sequence 125, App
29	173	5.4	419	9	US-09-880-505-125	Sequence 125, App
30	172.5	5.3	112	10	US-09-205-658-75	Sequence 75, Appl
31	172.5	5.3	112	10	US-09-844-353A-75	Sequence 75, Appl
32	149	4.6	107	10	US-09-205-658-76	Sequence 76, Appl
33	149	4.6	107	10	US-09-844-353A-76	Sequence 76, Appl
34	137.5	4.3	88	10	US-09-751-100B-62	Sequence 62, Appl
35	137	4.2	85	10	US-09-751-100B-59	Sequence 59, Appl
36	137	4.2	85	10	US-09-751-100B-60	Sequence 60, Appl
37	137	4.2	85	10	US-09-751-100B-61	Sequence 61, Appl
38	134.5	4.2	90	10	US-09-751-100B-49	Sequence 49, Appl
39	131.5	4.1	90	10	US-09-751-100B-48	Sequence 48, Appl
40	131	4.1	24	9	US-10-106-698-5238	Sequence 5238, Ap
41	131	4.1	85	10	US-09-751-100B-54	Sequence 54, Appl
42	131	4.1	85	10	US-09-751-100B-55	Sequence 55, Appl
43	131	4.1	85	10	US-09-751-100B-56	Sequence 56, Appl
44	131	4.1	85	10	US-09-751-100B-57	Sequence 57, Appl
45	131	4.1	85	10	US-09-751-100B-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-10-205-823-162
; Sequence 162, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-162

Query Match 24.4%; Score 789.5; DB 9; Length 690;
Best Local Similarity 35.1%; Pred. No. 5.3e-60;

```
  Matches  207;  Conservative   96;  Mismatches  219;  Indels   67;  Gaps   15;

Qy         30 QLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEIL-QMFGKMFFVFCQESGYDTILR 88
              ::|  : | |  :: :     |||   |      |:: :  |:  |  | |   : ||
Db        101 KIKESRKSLEREDFEKTIAEQAVAAGVPV------EVIKESLGEEVFKICYEED-ENILG 153

Qy         89 VLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCTDAEKGKGLILHYYSEREGLQ 143
              |:|  :::|| :   |     |   |      :  | | | |    | ::|:  :
Db        154 VVGGTLKDFLNSFSTLLKQSSHCQEAGKRGRLEDASILCLDKE-DDFLHVYYFFPKRTTS 212

Qy        144 DIVIGIIKTVAQQIHGTEIDMKVIQQRNEECDHTQ-----------FLIEEKESKEEDFY 192
               |: ||||  |  :: ||::: ::      | |              : :  | :|
Db        213 LILPGIIKAAAHVLYETEVEVSLM----PPCFHNDCSEFVNQPYLLYSVHMKSTKPS--- 265

Qy        193 EDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQ-PGNCSL 250
                    :  | | : | : ||| |||| :||:|: : | || | |:: :    |  :
Db        266 -----LSPSKPQSSLVIPTSLFCKTFPFHFMFDKDMTILQFGNGIRRLMNRRDFQGKPNF 320

Qy        251 LSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQM 310
               | :: | |: :| ||:: :|  ||:| :     |:|              : |||||
Db        321 EEYFEILTPKINQTFSGIMTMLNMQFVVRVRRWDNSVKK----------SSRVMDLKGQM 370

Qy        311 IYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQFREEYKLTQELE 370
               ||: |: :|||| || |  |:| | |||||||||:|:| ||:||:||| | :  | : |
Db        371 IYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQARAQDGLKKRLG 430

Qy        371 ILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTILFSGIVGF 430
               |   |:   :|||:|||||  || |: |  || :|   : | ||:: |||:||| ||||
Db        431 KLKATLEQAHQALEEEKKKTVDLLCSIFPCEVAQQLWQGQVVQAKKFSNVTMLFSDIVGF 490

Qy        431 NAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCI 490
               | ||: :     :::: :|| ||||||           ||||||:|| |    || :
Db        491 TAICSQCS----PLQVITMLNALYTRFDQQCGELD---VYKVETIGDAYCVAGGLHKESD 543

Qy        491 HHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNL 549
               ||  |  :|| |||:: :|    || ::: ||:|:| |  ||:| :|||||||| | |
Db        544 THAVQIALMALKMMELSDEVMSPHGEPIKMRIGLHSGSVFAGVVGVKMPRYCLFGNNVTL 603

Qy        550 TSRTETTGEKGKINVSEYTYRCL----------MSPENSDPQFHLEHRG 588
               :: |:     |||||  ||| |           | |   | |  | |
Db        604 ANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELPPNFPSEIPG 652


RESULT 2
US-09-952-213D-2
; Sequence 2, Application US/09952213D
; Publication No. US20030096240A1
; GENERAL INFORMATION:
;  APPLICANT: MURAD, FERID
;  APPLICANT:  SHARINA, IRAIDA G.
;  APPLICANT:  KRUMENACKER, J. S.
;  APPLICANT:  MARTIN, E.
;  TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC
;  FILE REFERENCE: UTSH:252US
;  CURRENT APPLICATION NUMBER: US/09/952,213D
;  CURRENT FILING DATE:  2002-08-16
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 691
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-952-213D-2

  Query Match           24.1%;  Score 778;  DB 9;  Length 691;
  Best Local Similarity   33.5%;  Pred. No. 5.3e-59;
  Matches  213;  Conservative  105;  Mismatches  224;  Indels   94;  Gaps   20;

Qy          9 LELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGEILQ 68
              | | : |       |: :| :: ::    | :|| :::     :|| : |  |
Db         89 LNLALQRTLAKHKIEENRKSSEKED----LEKIIAEEA-----IAAGAPVEALK-----D 134

Qy         69 MFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDAL-----HDHLATIYPGMRAPSFRCT 123
                |:  |  | |   : || |:|  :::|| :   |     |   |      :   | |
Db        135 SLGEELFKICYEED-EHILGVVGGTLKDFLNSFSTLLKQSSHCQEAERRGRLEDASILCL 193

Qy        124 DAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNEEC---DHTQFL 180
              | ::   | ::|:  :     :: ||||  |: :: : ::: ::      |  | |:|:
Db        194 DKDQ-DFLNVYYFFPKRTTALLLPGIIKAAARILYESHVEVSLM----PPCFRSDCTEFV 248

Qy        181 --------IEEKESKEEDFYEDLDRFEENGTQESRISPYT-FCKAFPFHIIFDRDLVVTQ 231
                     :  | :|                  | | : | : ||| |||| : |||| : |
Db        249 NQPYLLYSVHVKSTKPS--------LSPGKPQSSLVIPASLFCKTFPFHFMLDRDLAILQ 300

Qy        232 CGNAIYRVLPQLQ-PGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKL 290
               || | |:: :     |  :    | :: | |: :| ||:: :|  ||:| :     |:|
Db        301 LGNGIRRLVNKRDFQGKPNFEEFFEILTPKINQTFSGIMTMLNMQFVIRVRRWDNSVKK- 359

Qy        291 ECEDELTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATR 350
                          : |||||||: |: :|||| || |   |:| | |||||||||:|:| |
Db        360 ---------SSRVMDLKGQMIYIVESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNALR 410

Qy        351 DLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELRHKR 410
               |:||:||| | :  | : |  |   |:   :|||:|||:|  || |: |  || :|   :
Db        411 DVVLIGEQARAQDGLKKRLGKLKATLEHAHQALEEEKKRTVDLLCSIFPSEVAQQLWQGQ 470

Qy        411 PVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVY 470
               | ||::  ||:||| |||| | ||: :     :::: :|| ||||||           ||
Db        471 IVQAKKFSEVTMLFSDIVGFTAICSQCS----PLQVITMLNALYTRFDQQCGELD---VY 523

Qy        471 KVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTGEVV 529
              ||||:|| |    ||       ||  |  :|| |||:: :|    || ::: ||:|:| |
Db        524 KVETIGDAYCVAGGLHRESDTHAVQIALMALKMMELSNEVMSPHGEPIKMRIGLHSGSVF 583

Qy        530 TGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL----------MSPENSD 579
               ||:| :|||||||| | | :: |:     |||||  ||| |           | |
Db        584 AGVVGVKMPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFTPRSREELP 643

Qy        580 PQF--------HL----EHRGPVSMKGKKEPMQVWF 603
               | |        |.      |:|| |         : ||
Db        644 PNFPSDIPGICHFLDAYHHQGPNS--------KPWF 671


RESULT 3
US-09-819-249-2
; Sequence 2, Application US/09819249
; Patent No. US20010029019A1
; GENERAL INFORMATION:
;  APPLICANT: Waldman, Scott A.
;  APPLICANT:  Park, Jason
;  APPLICANT:  Schulz, Stephanie
;  TITLE OF INVENTION: Compositions And Methods For Identifying And Targeting Cancer
;  TITLE OF INVENTION:  Alimentary Canal Origin
;  FILE REFERENCE: TJU2412
;  CURRENT APPLICATION NUMBER: US/09/819,249
;  CURRENT FILING DATE:  2001-03-27
;  PRIOR APPLICATION NUMBER: 60/192,229
;  PRIOR FILING DATE: 2000-03-27
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 1073
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-819-249-2

  Query Match           12.6%;  Score 407.5;  DB 10;  Length 1073;
  Best Local Similarity   22.6%;  Pred. No. 2e-26;
  Matches  180;  Conservative  117;  Mismatches  274;  Indels  225;  Gaps   26;

Qy         12 LVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAAS-----KVLNLNAGEI 66
              ::|   |||    :| :  : |:   ::  :::|    | | |        || |: |
Db        245 VIIMCGGPEFLYKLKGDRAVAEDIVIILVDLFNDQYLEDNVTAPDYMKNVLVLTLSPGNS 304

Qy         67 L----------------------QMFGKMFFVFCQE-----------------SGYD 84
              |                       :|| |  :| :                  |||
```

```
Db        305 LLNSSFSRNLSPTKRDFRLAYLNGILVFGHMLKIFLENGENITTPKFAHAFRNLTFEGYD 364

Qy         85 ------------TILRVLGSNVREFLQNLDALHDHLATIYPGMRAPSFRCTDAE----- 126

Db        365 GPVTLDDWGDVDSTMVLLYTSVDTKKYKVLLTYDTHVNKTYPVDMSPTFTWKNSKLPNDI 424

Qy        127 KGKG---LILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQR------------ 170

Db        425 TGRGPQILMIAVFT----LTGAVVLLLLVALLMLRKYRKDYELRQKKWSHIPPENIFPLE 480

Qy        171 NEECDHTQFLIEE------------KESKEEDFYEDLDRFEENGTQESR--------IS 209

Db        481 TNETNHVSLKIDDDKRRDTIQRLRQCKYVKKRVILKDLKHNDGNFTEKQKIELNKLLQID 540

Qy        210 PYTFCKAFPFHIIFDRDLVVTQCGNAIYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGIL 269

Db        541 YYTLTKFYG----------TVKLDTMIFGVIEYCERG--SLREVLNDTISYPDGTFMDWE 588

Qy        270 SHINTVFVLRSKEGLLDVEKLECEDELTGTEISCLRLKGQMIYLPE--ADSIL------- 320

Db        589 FKISVLYDIAKGMSYLHSSKTEVHGRLKST--NCVVDSRMVVKITDFGCNSILPPKKDLW 646

Qy        321 ----FLCSPSVMNLDDLTRRGLYLSDIPLHDAT------RD------------------- 351

Db        647 TAPEHLRQANISQKGDVYSYGIIAQEIILRKETFYTLSCRDRNEKIFRVENSNGMKPFRP 706

Qy        352 ---LVLLGEQFREEYKLTQE---------------------------------LEILTD 374

Db        707 DLFLETAEEKELEVYLLVKNCWEEDPEKRPDFKKIETTLAKIFGLFHDQKNESYMDTLIR 766

Qy        375 RLQLTLRALED----------EKKKTDTLLYSVLPPSVANELRHKRPVPAKRYDNVTIL 423

Db        767 RLQLYSRNLEHLVEERTQLYKAERDRADRLNFMLLPRLVVKSLKEKGFVEPELYEEVTIY 826

Qy        424 FSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVS 483

Db        827 FSDIVGFTTIC-KYST---PMEVVDMLNDIYKSFDHIVDHHD---VYKVETIGDAYMVAS 879

Qy        484 GLPE-PCIHHARSICHLALDMMEIAGQVQVD---GESVQITIGIHTGEVVTGVIGQRMPR 539

Db        880 GLPKRNGNRHAIDIAKMALEILSFMGTFELEHLPGLPIWIRIGVHSGPCAAGVVGIKMPR 939

Qy        540 YCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLEHRGPVSMKGKKEPM 599

Db        940 YCLFGDTVNTASRMESTGLPLRIHVSGSTIAIL---KRTECQFLYEVRGETYLKGRGNET 996

Qy        600 QVWFLSRKNTGTEETK 615

Db        997 TYWL-----TGMKDQK 1007


RESULT 4
US-10-157-031-18
; Sequence 18, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
;  APPLICANT: Baranova, A. V.
;  APPLICANT:  Yankovsky, N. K.
;  APPLICANT:  Kozlov, A. P.
;  APPLICANT:  Lobashev, A. V.
;  APPLICANT:  Krukovskaya, L. L.
;  TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
;  FILE REFERENCE: 2760-103
;  CURRENT APPLICATION NUMBER: US/10/157,031
;  CURRENT FILING DATE:  2002-05-30
;  NUMBER OF SEQ ID NOS: 415
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 1073
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-157-031-18

  Query Match           12.6%;  Score 405.5;  DB 9;  Length 1073;
  Best Local Similarity  37.2%;  Pred. No. 3e-26;
  Matches  103;  Conservative   44;  Mismatches   95;  Indels   35;  Gaps    9;

Qy        354 LLGEQFREEYKLTQELEILTDRLQLTLRALED----------EKKKTDTLLYSVLPPSV 402

Db        751 LFHDQKNESY-----MDTLIRRLQLYSRNLEHLVEERTQLYKAERDRADRLNFMLLPRLV 805

Qy        403 ANELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTD 462

Db        806 VKSLKEKGFVEPELYEEVTIYFSDIVGFTTIC-KYST---PMEVVDMLNDIYKSFDHIVD 861

Qy        463 SRKNPFVYKVETVGDKYMTVSGLPE-PCIHHARSICHLALDMMEIAGQVQVD---GESVQ 518

Db        862 HHD---VYKVETIGDAYMVASGLPKRNGNRHAIDIAKMALEILSFMGTFELEHLPGLPIW 918

Qy        519 ITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENS 578

Db        919 IRIGVHSGPCAAGVVGIKMPRYCLFGDTVNTASRMESTGLPLRIHVSGSTIAIL---KRT 975

Qy        579 DPQFHLEHRGPVSMKGKKEPMQVWFLSRKNTGTEETK 615

Db        976 ECQFLYEVRGETYLKGRGNETTYWL-----TGMKDQK 1007


RESULT 5
US-10-071-223-2
; Sequence 2, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
;  APPLICANT: Storm, Daniel R.
;  APPLICANT:  Hacker, Beth
;  APPLICANT:  Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  APPLICANT:  University of Washington
;  TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE: 44481-5029-02-US
;  CURRENT APPLICATION NUMBER: US/10/071,223
;  CURRENT FILING DATE:  2002-02-11
;  PRIOR APPLICATION NUMBER: 09/473,717
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: PCT/US98/13541
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/098,559
;  PRIOR FILING DATE: 1997-07-01
;  PRIOR APPLICATION NUMBER: 08/886,440
;  PRIOR FILING DATE: 1997-07-01
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1294
;   TYPE: PRT
;   ORGANISM: human type IX adenylyl cyclase
US-10-071-223-2

  Query Match           9.0%;  Score 291.5;  DB 12;  Length 1294;
  Best Local Similarity  31.2%;  Pred. No. 3.5e-16;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL----------RHK---------- 409

Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469

Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEETK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528

Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
```

```
                : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 6
US-09-751-100B-2
; Sequence 2, Application US/09751100B
; Patent No. US20020142436A1
; GENERAL INFORMATION:
;  APPLICANT: Medical Research Council
;  TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor
;  FILE REFERENCE: P27948A
;  CURRENT APPLICATION NUMBER: US/09/751,100B
;  CURRENT FILING DATE:  2000-12-28
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: Mouse
US-09-751-100B-2

  Query Match           9.0%;  Score 291.5;  DB 10;  Length 1353;
  Best Local Similarity   31.2%;  Pred. No. 3.7e-16;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL-------------RHK-------------------- 409
              || ||   : :::||:|  :|:::|              ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEQTK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
               |: |:|| |  |:| ||| |  :| : | |::  |  |  | | : :|:||| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 7
US-10-071-223-3
; Sequence 3, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
;  APPLICANT: Storm, Daniel R.
;  APPLICANT:  Hacker, Beth
;  APPLICANT:  Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  APPLICANT:  University of Washington
;  TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE: 44481-5029-02-US
;  CURRENT APPLICATION NUMBER: US/10/071,223
;  CURRENT FILING DATE:  2002-02-11
;  PRIOR APPLICATION NUMBER: 09/473,717
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: PCT/US98/13541
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/098,559
;  PRIOR FILING DATE: 1997-07-01
;  PRIOR APPLICATION NUMBER: 08/886,440
;  PRIOR FILING DATE: 1997-07-01
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: murine type IX adenylyl cyclase
US-10-071-223-3

  Query Match           9.0%;  Score 291.5;  DB 12;  Length 1353;
  Best Local Similarity   31.2%;  Pred. No. 3.7e-16;
  Matches   74;  Conservative   44;  Mismatches   72;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL-------------RHK-------------------- 409
              || ||   : :::||:|  :|:::|              ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEQTK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
               |: |:|| |  |:| ||| |  :| : | |::  |  |  | | : :|:||| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTGTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 8
US-10-282-942-2
; Sequence 2, Application US/10282942
; Publication No. US20030087295A1
; GENERAL INFORMATION:
;  APPLICANT: Silos-Santiago, Inmaculada
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
;  TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF PAIN DISORDERS USING 9805
;  FILE REFERENCE: MPI01-240P1RM
;  CURRENT APPLICATION NUMBER: US/10/282,942
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: 60/335,047
;  PRIOR FILING DATE: 2001-10-31
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1080
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-282-942-2

  Query Match           8.9%;  Score 289;  DB 9;  Length 1080;
  Best Local Similarity   26.3%;  Pred. No. 4.4e-16;
  Matches   94;  Conservative   66;  Mismatches  109;  Indels   88;  Gaps   12;

Qy        236 IYRVLPQLQPGNCSLLSVFSLVRPHIDISFHGILSHINTVFVLRSKEGLLDVEKLECEDE 295
              :| :||    |  ::                | :|  : : || |  |
Db        136 VYTLLPFSMRGAVAV---------------GAVSTASHLLVLGSLMG----------- 167

Qy        296 LTGTEISCLRLKGQMIYLPEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLL 355
                |     :|:  |::    |::::|||     ||     :        : ||:|||
Db        168 --GFTTPSVRVGLQLL----ANAVIFLCG----NLTGAFHKH------QMQDASRDL--- 208

Qy        356 GEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVA----------N 404
                    |  :: :  |     | |  ||:: : || ||||  ::
Db        209 ---------FTYTVKCIQIR-----RKLRIEKRQQENLLLSVLPAHISMGMKLAIIERLK 254

Qy        405 ELRHKRPVP--------AKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTR 456
              |   :| :|          ||: ||:||:: ||||    |  :  |    :| :||:|: :
Db        255 EHGDRRCMPDNNFHSLYVKRHQNVSILYADIVGFTQLASDCSPKE----LVVVLNELFGK 310

Qy        457 FDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDGE 515
              || :  : :     ::: :|| |  |||||     |||:   : ||| :   || :  |
Db        311 FDQIAKANE---CMRIKILGDCYYCVSGLPVSLPTHARNCVKMGLDMCQAIKQVREATGV 367

Qy        516 SVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCL 572
               : : :|||:| |: |||| |  :| :: : |:| :| |  |  |:::::| | : |
Db        368 DINMRVGIHSGNVLCGVIGLRKWQYDVWSHDVSLANRMEAAGVPGRVHITEATLKHL 424


RESULT 9
```

```
US-09-751-100B-99
; Sequence 99, Application US/09751100B
; Patent No. US20020142436A1
; GENERAL INFORMATION:
;  APPLICANT: Medical Research Council
;  TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor
;  FILE REFERENCE: P27948A
;  CURRENT APPLICATION NUMBER: US/09/751,100B
;  CURRENT FILING DATE:  2000-12-28
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
;   LENGTH: 1353
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-751-100B-99

  Query Match             8.8%;  Score 283.5;  DB 10;  Length 1353;
  Best Local Similarity   30.8%;  Pred. No. 1.9e-15;
  Matches   73;  Conservative   44;  Mismatches   73;  Indels   47;  Gaps    6;

Qy        383 LEDEKKKTDTLLYSVLPPSVANEL----------RHK------------------- 409
              || ||   : :::|||:|  :|::|           ||
Db        323 LEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF 382

Qy        410 RPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFV 469
              ||   :: : |:|||: ||||    :| :: : |  :| |||||: ||| | :  |
Db        383 RPFKMQQIEEVSILFADIVGF----TKMSANKSAHALVGLLNDLFGRFDRLCEETK---C 435

Qy        470 YKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQ-VQVDGESVQITIGIHTGEV 528
               |: |:|| |  |:| ||| || : | |::  |  |  | | : :|:|| |
Db        436 EKISTLGDCYYCVAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKKEMVNMRVGVHTRTV 495

Qy        529 VTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              : |::| |  :: :: | ||| :  |  |  ||:::|| | : |      | :: :|
Db        496 LCGILGMRRFKFDVWSNDVNLANLMEQLGVAGKVHISEATAKYL------DDRYEME 546


RESULT 10
US-10-201-000-2
; Sequence 2, Application US/10201000
; Publication No. US20020187540A1
; GENERAL INFORMATION:
;  APPLICANT: Tomlinson, James E.
;  APPLICANT:  COR Therapeutics, Inc.
;  TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE: 44481-5028-01-US
;  CURRENT APPLICATION NUMBER: US/10/201,000
;  CURRENT FILING DATE:  2002-07-24
;  PRIOR APPLICATION NUMBER: US/09/474,076
;  PRIOR FILING DATE: 1999-12-12
;  PRIOR APPLICATION NUMBER: PCT/US98/13694
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/070,904
;  PRIOR FILING DATE: 1997-07-01
;  PRIOR APPLICATION NUMBER: 08/886,550
;  PRIOR FILING DATE: 1997-07-01
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1168
;   TYPE: PRT
;   ORGANISM: human type VI adenylyl cyclase
US-10-201-000-2

  Query Match             8.7%;  Score 282.5;  DB 9;  Length 1168;
  Best Local Similarity   26.6%;  Pred. No. 1.8e-15;
  Matches   94;  Conservative   61;  Mismatches  126;  Indels   73;  Gaps   12;

Qy        296 LTGTEISCLRL-------KGQMIYLPE--ADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346
              |:|  :| | |       :|      :  |: :||||:
Db        262 LSGLGLSTLHLILAWQLNRGDAFLWKQLGANVLLFLCT---------------------- 299

Qy        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406
              ::: :   :  |    | :     :|  |  |: | :: : || ||||  || |:
Db        300 ----NVIGICTHYPAEVSQRQAFQETRGYIQARLH-LQHENRQQERLLLSVLPQHVAMEM 354

Qy        407 R------------HKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLY 454
              :            ||  :  :::|||:|||: | || :  |:      | ::|  ||:|:
Db        355 KEDINTKKEDMMFHK--IYIQKHDNVSILFADIEGFTSLASQCT----AQELVMTLNELF 408

Qy        455 TRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVD 513
               ||| |          ::: :|| |  ||||||    ||     : :||:|    | :|
Db        409 ARFDKLAAEN---HCLRIKILGDCYYCVSGLPEARADHAHCCVEMGVDMIEAISLVREVT 465

Qy        514 GESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLM 573
              | :| : :|||:| |  ||:| |  :: :: | | | :  |  |  |:|:::  | : |
Db        466 GVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHMEAGGRAGRIHITRATLQYLN 525

Qy        574 SPENSDP-----------QFHLEH---RGPVSMKGKKEPMQVWFLSRKNTGTEE 613
                  :|           : |:|     |  | | |:|   :  | |    : |
Db        526 GDYEVEPGRGGERNAYLKEQHIETFLILG-ASQKRKEEKAMLAKLQRTRANSME 578


RESULT 11
US-09-750-240-11
; Sequence 11, Application US/09750240
; Patent No. US20020103147A1
; GENERAL INFORMATION:
;  APPLICANT: Hammon, H. K.
;  APPLICANT:  Insel, P. A.
;  APPLICANT:  Ping, P.
;  APPLICANT:  Post, S. R.
;  APPLICANT:  Gao, M.
;  TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
;  TITLE OF INVENTION:  FAILURE
;  FILE REFERENCE: 220002056723
;  CURRENT APPLICATION NUMBER: US/09/750,240
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 09/472,667
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: US 09/008,097
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: US 08/924,757
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: US 60/048,933
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: US 08/708,661
;  PRIOR FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 1168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-750-240-11

  Query Match             8.7%;  Score 282.5;  DB 10;  Length 1168;
  Best Local Similarity   26.6%;  Pred. No. 1.8e-15;
  Matches   94;  Conservative   61;  Mismatches  126;  Indels   73;  Gaps   12;

Qy        296 LTGTEISCLRL-------KGQMIYLPE--ADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346
              |:|  :| | |       :|      :  |: :||||:
Db        262 LSGLGLSTLHLILAWQLNRGDAFLWKQLGANVLLFLCT---------------------- 299

Qy        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406
              ::: :   :  |    | :     :|  |  |: | :: : || ||||  || |:
Db        300 ----NVIGICTHYPAEVSQRQAFQETRGYIQARLH-LQHENRQQERLLLSVLPQHVAMEM 354

Qy        407 R------------HKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLY 454
              :            ||  :  :::|||:|||: | || :  |:      | ::|  ||:|:
Db        355 KEDINTKKEDMMFHK--IYIQKHDNVSILFADIEGFTSLASQCT----AQELVMTLNELF 408

Qy        455 TRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVD 513
               ||| |          ::: :|| |  ||||||    ||     : :||:|    | :|
```

```
Db        409 ARFDKLAAEN---HCLRIKILGDCYYCVSGLPEARADHAHCCVEMGVDMIEAISLVREVT 465

Qy        514 GESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLM 573

Db        466 GVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHMEAGGRAGRIHITRATLQYLN 525

Qy        574 SPENSDP-----------QFHLEH---RGPVSMKGKKEPMQVWFLSRKNTGTEE 613

Db        526 GDYEVEPGRGGERNAYLKEQHIETFLILG-ASQKRKEEKAMLAKLQRTRANSME 578


RESULT 12
US-09-750-240-4
; Sequence 4, Application US/09750240
; Patent No. US20020103147A1
; GENERAL INFORMATION:
;  APPLICANT: Hammon, H. K.
;  APPLICANT:  Insel, P. A.
;  APPLICANT:  Ping, P.
;  APPLICANT:  Post, S. R.
;  APPLICANT:  Gao, M.
;  TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
;  TITLE OF INVENTION:  FAILURE
;  FILE REFERENCE: 220002056723
;  CURRENT APPLICATION NUMBER: US/09/750,240
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 09/472,667
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: US 09/008,097
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: US 08/924,757
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: US 60/048,933
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: US 08/708,661
;  PRIOR FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 604
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-750-240-4

  Query Match             8.6%;  Score 278.5;  DB 10;  Length 604;
  Best Local Similarity   27.1%;  Pred. No. 1.5e-15;
  Matches   83;  Conservative   56;  Mismatches  110;  Indels   57;  Gaps    9;

Qy        296 LTGTEISCLRL-------KGQMIYLPE--ADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346

Db         51 LSGLGLSTLHLILAWQLNRGDAFLWKQLGANVLLFLCT---------------------- 88

Qy        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406

Db         89 ----NVISICTHYPAEVSQRQAFQETRSYIQARLH-LQHENRQQERLLLSVLPQHVAMEM 143

Qy        407 R-----------HKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYT 455

Db        144 KEDINTKKEDMFHK--IYIQKHDNVSILFADIEGFTSLASQCT----AQELVMTLNELFA 197

Qy        456 RFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDG 514

Db        198 RFDKLAAEN---HCLRIKILGDCYYCVSGLPEARADHAHCCVEMGVDMIEAISLVREVTG 254

Qy        515 ESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMS 574

Db        255 VNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHMEAGSRAGRIHITRATLQYLNG 314

Qy        575 PENSDP 580

Db        315 DYEVEP 320


RESULT 13
US-09-750-240-6
; Sequence 6, Application US/09750240
; Patent No. US20020103147A1
; GENERAL INFORMATION:
;  APPLICANT: Hammon, H. K.
;  APPLICANT:  Insel, P. A.
;  APPLICANT:  Ping, P.
;  APPLICANT:  Post, S. R.
;  APPLICANT:  Gao, M.
;  TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
;  TITLE OF INVENTION:  FAILURE
;  FILE REFERENCE: 220002056723
;  CURRENT APPLICATION NUMBER: US/09/750,240
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: US 09/472,667
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: US 09/008,097
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: US 08/924,757
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: US 60/048,933
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: US 08/708,661
;  PRIOR FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 1167
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-750-240-6

  Query Match             8.6%;  Score 278.5;  DB 10;  Length 1167;
  Best Local Similarity   27.1%;  Pred. No. 4e-15;
  Matches   83;  Conservative   56;  Mismatches  110;  Indels   57;  Gaps    9;

Qy        296 LTGTEISCLRL-------KGQMIYLPE--ADSILFLCSPSVMNLDDLTRRGLYLSDIPLH 346

Db        262 LSGLGLSTLHLILAWQLNRGDAFLWKQLGANVLLFLCT---------------------- 299

Qy        347 DATRDLVLLGEQFREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANEL 406

Db        300 ----NVISICTHYPAEVSQRQAFQETRSYIQARLH-LQHENRQQERLLLSVLPQHVAMEM 354

Qy        407 R-----------HKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYT 455

Db        355 KEDINTKKEDMFHK--IYIQKHDNVSILFADIEGFTSLASQCT----AQELVMTLNELFA 408

Qy        456 RFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDG 514

Db        409 RFDKLAAEN---HCLRIKILGDCYYCVSGLPEARADHAHCCVEMGVDMIEAISLVREVTG 465

Qy        515 ESVQITIGIHTGEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMS 574

Db        466 VNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHMEAGSRAGRIHITRATLQYLNG 525

Qy        575 PENSDP 580

Db        526 DYEVEP 531


RESULT 14
US-10-175-158-2
; Sequence 2, Application US/10175158
; Publication No. US20030008371A1
; GENERAL INFORMATION:
;  APPLICANT: Tomlinson, James
;  APPLICANT:  Cor Therapeutics, Inc.
;  TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
;  TITLE OF INVENTION:  CYCLASE
;  FILE REFERENCE: 44481-5027-01-US
;  CURRENT APPLICATION NUMBER: US/10/175,158
```

; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US/09/473,716
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PCT/US98/13540
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/070,901
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 08/886,362
; PRIOR FILING DATE: 1997-07-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: human type V adenylyl cyclase
US-10-175-158-2

Query Match 8.6%; Score 278; DB 9; Length 1261;
Best Local Similarity 28.8%; Pred. No. 5e-15;
Matches 80; Conservative 57; Mismatches 99; Indels 42; Gaps 9;

```
Qy        360 REEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVANELR---------- 407
              |: :: |:|        :|  | : : | :: : || ||||  || |::
Db        405 RQAFQETREC------IQARLHS-QRENQQQERLLLSVLPRHVAMEMKADINAKQEDMMF 457

Qy        408 HKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNP 467
              ||  :  :::|||:|||: | || :  |:      | ::|  ||:|: ||| |
Db        458 HK--IYIQKHDNVSILFADIEGFTSLASQCT----AQELVMTLNELFARFDKLAAEN--- 508

Qy        468 FVYKVETVGDKYMTVSGLPEPCIHHARSICHLALDMMEIAGQV-QVDGESVQITIGIHTG 526
                 ::: :|| |  ||||||    ||     : :||:|    | :| | :| : :|||:|
Db        509 HCLRIKILGDCYYCVSGLPEARADHAHCCVEMGMDMIEAISLVREVTGVNVNMRVGIHSG 568

Qy        527 EVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDP------ 580
               |  ||:| |  :: :: | |: | :  |  |: |:|:::: |   |      :|
Db        569 RVHCGVLGLRKWQFDVWSNDVTLANHMEAGGKAGRIHITKATLNYLNGDYEVEPGCGGER 628

Qy        581 -QFHLEHRGPV------SMKGKKEPMQVWFLSRKNTGT 611
               :  ||          : | |:|   :  ::|: | :
Db        629 NAYLKEHSIETFLILRCTQKRKEEKAMIAKMNRQRTNS 666
```

RESULT 15
US-09-989-442-95
; Sequence 95, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ08
; CURRENT APPLICATION NUMBER: US/09/989,442
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Query Match 8.3%; Score 269.5; DB 9; Length 257;
Best Local Similarity 29.6%; Pred. No. 2.4e-15;
Matches 77; Conservative 58; Mismatches 100; Indels 25; Gaps 9;

```
Qy        361 EEYKLTQELEILTDRLQLT--LRALEDEKKKTDTLLYSVLPPSV-ANELRHKRPVPAKRY 417
              :: : |  |: |  ||| |  :  :|: :     ||:::||  | |: |  :|      |
Db          2 QQVESTARLDFLW-RLQATEEIEEMEELQAYNRRLLHNILPKDVAAHFLARERRNDELYY 60

Qy        418 DN---VTILFSGIVGFNAFCSKHASGEGAMKIVNLLNDLYTRFDTLTDSRKNPFVYKVET 474
               :   | ::|: |  |: |  :  :    :: : |||:;   || :    :   : |::|
Db         61 QSCECVAVMFASIANFSEFYVELEANNEGVECLRLLNEIIADFDEIISEDRFRQLEKIKT 120

Qy        475 VGDKYMTVSGLPEPCIH-----HARSICHLALDMMEIAGQVQVDGE----SVQITIGIHT 525
              :|  ||  ||| :         | :::   |: :|:   |::   |    : |: ||::
Db        121 IGSTYMAASGLNDSTYDKVGKTHIKALADFAMKLMD---QMKYINEHSFNNFQMKIGLNI 177

Qy        526 GEVVTGVIGQRMPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMSPENSDPQFHLE 585
              | || |||| | |:| ::|||||: || ::||   :| |:   |: | :      : ||
Db        178 GPVVAGVIGARKPQYDIWGNTVNVASRMDSTGVPDRIQVTTDMYQVLAA-----NTYQLE 232

Qy        586 HRGPVSMKGKKEPMQVWFLS 605
              |  | :||| | |  :||:
Db        233 CRXVVKVKGKGE-MMTYFLN 251
```

Search completed: June 27, 2003, 13:17:13
Job time : 34.6287 secs